

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: D. Schreiner

Searcher Phone #: 308-4242

Searcher Location: CMI 12618

Date Searcher Picked Up: _____

Date Completed: 5/1

Searcher Prep & Review Time: 9

Clerical Prep Time: _____

Online Time: 82

Type of Search

NA Sequence (#) 2

AA Sequence (#) 3

Structure (#) _____

Bibliographic _____

Litigation _____

Fulltext _____

Patent Family _____

Other _____

Vendors and cost where applicable

STN: _____

Dialox: _____

Questel/Orbit: _____

Dr-Link: _____

Lexis/Nexis: _____

Sequence Systems: Compugen Ig Genet FastDB

WWW/Internet: _____

Other (specify): _____

The first part of the paper discusses the importance of understanding the cultural context of the research. It highlights the need for researchers to be sensitive to the values and beliefs of the communities they are studying. This is particularly important in the field of education, where cultural differences can significantly impact learning outcomes.

The second part of the paper focuses on the methodology used in the study. It describes the process of selecting participants, collecting data, and analyzing the results. The authors emphasize the importance of using a mixed-methods approach to gain a comprehensive understanding of the research topic.

The third part of the paper presents the findings of the study. It discusses the results of the quantitative data analysis and the insights gained from the qualitative interviews. The authors conclude that there are significant differences in learning outcomes between the two groups, and these differences can be attributed to cultural factors.

The final part of the paper discusses the implications of the findings for future research and practice. It suggests that educators should be aware of the cultural context of their students and tailor their teaching methods accordingly. Additionally, it recommends that future research should continue to explore the relationship between culture and learning outcomes.

XX New isolated LAGE-1 tumour associated nucleic acids - used to
 PT develop products for the diagnosis and treatment of LAGE-1
 PT associated disorders, particularly tumours
 XX
 PS claim 1; Page 50 52, 73pp; English.
 XX
 CC The present sequence encodes LAGE-1 tumour associated protein (TAP).
 CC The present invention also describes: (1) a method for treating a
 CC subject with a disorder characterised by expression of a LAGE-1 nucleic
 CC acid molecule or an expression product, comprising administering to the
 CC subject autologous cytolytic T cells to ameliorate the disorder, where
 CC the cytolytic T cells are specific for complexes of an HLA molecule and
 CC a LAGE-1 TAP or an immunogenic fragment, (2) a method for treating a
 CC subject with a disorder characterised by expression of a LAGE-1 nucleic
 CC acid molecule or an expression product, comprising administering a
 CC LAGE-1 TAP or an immunogenic fragment to ameliorate the disorder; and
 CC (3) a method for selectively enriching a population of T cells with
 CC cytolytic T cells specific for a LAGE-1 TAP comprising contacting an
 CC isolated population of T cells with an agent presenting a complex of a
 CC LAGE TAP or an immunogenic fragment and a HLA presenting molecule to
 CC selectively enrich the isolated population of T cells with the cytolytic
 CC T cells. The methods and products from the present invention can be used
 CC for the diagnosis and treatment of LAGE-1 associated disorders,
 CC particularly tumours.
 XX
 SS Sequence 1002 BP: 170 A: 302 C: 336 G: 185 T: 6 other:

Query Match 100 0%; Score 1002; DB 19; Length 1002;
 Best Local Similarity 100 0%; Prod. No. 1s 215;
 Matches 1002; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tctgctccacatctctgtgggcccctgacutctctctgagacggcgagagggctccgg 60
 Db 1 tctgctccacatctctgtgggcccctgacutctctctgagacggcgagagggctccgg 60
 QY 61 atccatgag 120
 Db 61 atccatgag 120
 QY 121 agacccatgag 180
 Db 121 agacccatgag 180
 QY 181 cag 240
 Db 181 cag 240
 QY 241 cag 300
 Db 241 cag 300
 QY 301 ggcacag 360
 Db 301 ggcacag 360
 QY 361 catcaaaag 420
 Db 361 catcaaaag 420
 QY 421 ag 480
 Db 421 ag 480
 QY 481 cag 540
 Db 481 cag 540
 QY 541 cag 600
 Db 541 cag 600

QY 601 accatgag 660
 Db 601 accatgag 660
 QY 661 cgcctttaaataatgatt 720
 Db 661 cgcctttaaataatgatt 720
 QY 721 gccaaatgag 780
 Db 721 gccaaatgag 780
 QY 781 cgcagatgatt 840
 Db 781 cgcagatgatt 840
 QY 841 agcctacag 900
 Db 841 agcctacag 900
 QY 901 gccagttcatt 960
 Db 901 gccagttcatt 960
 QY 961 tttctgag 1002
 Db 961 tttctgag 1002

RESULT 2
 ID D00151 standard; cDNA; 993 BP.
 XX AC D00151;
 XX DT 31-JUL-2000 (first entry)
 XX DE Human LAGE-1 unsplined variant, LAGE-11 encoding cDNA.
 XX KW LAGE-11; CAMEL, CTL-recognised Antigen on MELANOMA; human; cancer; CTL;
 KW cytotoxic T lymphocyte; tumour-associated antigen, NY-ESO-1, anticancer;
 KW melanoma; immunotherapy; immune response, ss.
 XX OS Homo sapiens.
 XX Key Location/Qualifiers
 FH 1..55
 FT 5'UTR /*tag= a
 FT CDS 56..688
 FT /*tag= b
 FT /product= "Human LAGE-11 protein"
 FT 689..993
 FT 3'UTR /note= "Unspliced variant of LAGE-1"
 FT /*tag= c
 XX PN W0200023584-A1.
 XX PD 27 APR-2000.
 XX PF 15-OCT-1995; 55WU-EP07832.
 XX PK 16-OCT-1998; 98EP-0119583.
 XX PA (BOEH) BOHRINGER INGELHEIM INT GMBH.
 PA (UYHO-) UNIV HOSPITAL LEIDEN.
 XX PI Schrier PL, Aarnoudse CA, Heider K, Klade C;
 XX WP1: 2000-339685/29.
 XX DR 3: FSLB, Y70861.
 PT Tumor-associated antigen useful for cancer immunotherapy is encoded by
 the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -

with the NY-ESO-1, a melanoma specific tumour antigen. The tumour-associated antigen displayed on melanoma cells is recognised by cytotoxic T lymphocytes. This sequence has antitumour activity (AME) (over antigen and immunogenic peptides derived from it are useful for cancer immunotherapy. They have the potential to induce an immune response, by eliciting a cell response. The DNA molecule is used for the construction of recombinant or fusion proteins.

Sequence 767 BP; 138 A; 234 C; 252 G; 143 T, 0 other,

Query Match 51.48; Score 515; DB 21; Length 767;
Best Local Similarity 76.48; Prod. No. 1.2e 106,
Matches 767. Concentrative 0. Mismatch 0. Total 0.0

| | | | |
|----|-----|--|-------|
| QY | 12 | atctctgttaacccctgaacctctctctctgaagccggccagagctctccgagaccatgcaagg | 71 |
| DB | 1 | | |
| QY | 72 | ccaaaggtccagggacagggggttcgaacggcgatgctcaatgagccaggaagaccctgaca | 141 |
| DB | 61 | | |
| QY | 132 | ttctctaatgcccgaatggaatgctatagggccacagadagagctgactgctcagagcgagca | 191 |
| DB | 121 | | |
| QY | 192 | gaattcccccgtgag | 251 |
| DB | 181 | | |
| QY | 252 | gtccctcatgaggttgaggtctctgctgcaatgagagagagagagagagagagagagag | 311 |
| DB | 241 | | |
| QY | 312 | cagacagagccctactcgaattacacatcacatcacccttctctcatcaccctggaagcag | 371 |
| DB | 301 | | |
| QY | 372 | aactatgtccacagatctctgtcccgatagccagacctctcccgacacacagagcgatctc | 431 |
| DB | 361 | | |
| QY | 432 | ctatgaatctccacctcccgagagctctctctctatctgagttctgagacagagacagat | 491 |
| DB | 421 | | |
| QY | 492 | aagagctgagcgatagagatgagatgagatgagatgagatgagatgagatgagatgagatg | 551 |
| DB | 458 | ----- | ----- |
| QY | 552 | atataagctatataatctcagacacacacacacacacacacacacacacacacacacac | 611 |
| DB | 458 | ----- | ----- |
| QY | 612 | caagagcgag | 671 |
| DB | 458 | ----- | ----- |
| QY | 672 | tattatctcttgcacctcaccatttagagagagagagagagagagagagagagagagagag | 731 |
| DB | 458 | ----- | ----- |
| QY | 732 | gctctgacagctctctatgctcagagagagagagagagagagagagagagagagagagag | 791 |
| DB | 492 | | |
| QY | 792 | ctccccctatttttgcctcagagctccctcagagagagagagagagagagagagagagag | 851 |
| DB | 552 | | |
| QY | 852 | ctctctaatctatgacctccctccctcagagagagagagagagagagagagagagagagag | 911 |
| DB | 611 | | |


```

Db 121 ttcctgatgcaccaggggcaatgctggggccagagagagggctgtccacagggcgca 180
QY 192 gaggcccccagagagagagagagagagagagagagagagagagagagagagagagag 251
Db 181 gaattcccccagagagagagagagagagagagagagagagagagagagagagagag 246
QY 252 atcccatcagagagagagagagagagagagagagagagagagagagagagagagag 311
Db 243 atcccatcagagagagagagagagagagagagagagagagagagagagagagagag 401
QY 312 cggagagagagagagagagagagagagagagagagagagagagagagagagagagag 371
Db 301 cggagagagagagagagagagagagagagagagagagagagagagagagagagagag 360
QY 372 agctatcagagagagagagagagagagagagagagagagagagagagagagagag 431
Db 361 agctatcagagagagagagagagagagagagagagagagagagagagagagagag 420
QY 432 taagagagagagagagagagagagagagagagagagagagagagagagagagagag 491
Db 421 taagagagagagagagagagagagagagagagagagagagagagagagagagagag 457
QY 492 aaagagagagagagagagagagagagagagagagagagagagagagagagagagag 551
Db 458 ----- 457
QY 552 agaaagagagagagagagagagagagagagagagagagagagagagagagagagag 611
Db 458 ----- 457
QY 612 caggcccccagagagagagagagagagagagagagagagagagagagagagagagag 671
Db 458 ----- 457
QY 672 taatttttctgcccctccatctttagccagagagagagagagagagagagagagag 731
Db 458 ----- 457
QY 732 atcccatcagagagagagagagagagagagagagagagagagagagagagagagag 791
Db 492 ctctccatcagagagagagagagagagagagagagagagagagagagagagagagag 551
QY 792 ctcccatcagagagagagagagagagagagagagagagagagagagagagagagag 851
Db 552 ctcccatcagagagagagagagagagagagagagagagagagagagagagagagag 911
QY 852 ctcccatcagagagagagagagagagagagagagagagagagagagagagagagag 911
Db 612 ctcccatcagagagagagagagagagagagagagagagagagagagagagagagag 671
QY 912 atcccatcagagagagagagagagagagagagagagagagagagagagagagagag 971
Db 672 atcccatcagagagagagagagagagagagagagagagagagagagagagagagag 731
QY 972 atcccatcagagagagagagagagagagagagagagagagagagagagagagagag 988
Db 732 atcccatcagagagagagagagagagagagagagagagagagagagagagagagag 748

```

RESULT 8

D00152
ID D00152 standard; cDNA; 752 BP.

XX AC D00152;

XX AC D00152;

DI 31-JUL-2000 (first entry)

DE Human tumour antigen, NY-ESO-1 cDNA.

KW NY ESO-1; CAMEL, CTL recognised Antigen on MELANOMA; human; cancer; CTL;
KW cytotoxic T lymphocyte; tumour-associated antigen; LAGE-1, anti-tumour,
KW melanoma, immunotherapy; immune response; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT 5'UTR 1..53

FT /tag= a

FT CDS 54..596

FT /tag= b

FT /product= "Human NY-ESO-1 protein"

FT /note= "Derived from open reading frame (ORF)-1"

FT 3'UTR 597..752

FT /tag= c

FT 5'UTR 1..93

FT /tag= d

FT /note= "5' UTR of ORF-2"

FT CDS 94..270

FT /tag= e

FT /product= "Human NY ESO 1 short variant protein"

FT /note= "Derived from open reading frame (ORF)-2"

FT 3'UTR 271..752

FT /tag= f

FT /note= "3' UTR of ORF 2"

XX W0200002584-A1.

XX 27-APR-2000.

XX 15-OCT-1999; 95WU-BI-07832.

XX 16-OCT-1998; 98EP-0119583.

XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.

XX (UYHO-) UNIV HOSPITAL LEIDEN.

XX Schrier PI, Aarnoudse CA, Heider K, Klade C;

XX WP1: 2000-339685/29.

XX E-PSUB; Y/0862, Y/0863.

XX Tumor associated antigen useful for cancer immunotherapy, is encoded by

XX the open reading frame of LAGE 1 (a tumor-specific antigen) cDNA -

XX Example 1; Page 61-62; 73pp; English.

XX The present sequence is the cDNA encoding human NY-ESO-1, a tumour

XX antigen, identified by screening a cDNA library of an esophagus

XX carcinoma. It has two different reading frames and the proteins encoded

XX contain epitopes of tumour specific proteins. NY-ESO-1 is expressed in

XX different tumour types, but not in healthy tissues except in testis. It

XX also shows homology with the CAMEL (Cytotoxic T lymphocytes (CTL)-

XX recognised Antigen on MELANOMA) cDNA, a tumour-associated antigen. The

XX cytotoxic T lymphocytes. This sequence has anticancer activity. CAMEL

XX tumour antigen and immunogenic peptides derived from it are useful for

XX cancer immunotherapy. They have the potential to induce an immune

XX response, by eliciting a CTL response. The DNA molecule is used for the

XX construction of recombinant or fusion proteins

XX Sequence 752 BP, 126 A, 236 C, 256 G, 149 T; 0 other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX


```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

om nucleic     nucleic search, using sw model

Run on:      May 1, 2001, 00:37:47 ; Search time 2512.86 Seconds
              (without alignments)
              580,645 Million cell updates/sec

Title:
Ref test score:  98.09 - 44 - 829A - 4
Sequence: 1 tctgtctcgtatctctgttq.....ctacgtcttcaaaaaaaa 1002

Scoring table:  IDENTITY_RSCF
Gapop 10.0 , Gapext 1.0

Searched: 128425 swps, 7374929652 residues

Total number of hits satisfying chosen parameters: 2566470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database:
GenBank: *
1:  qb_ba1.*
2:  qb_ba2.*
3:  qb_ba3.*
4:  qb_ba11.*
5:  qb_ba21.*
6:  qb_ba31.*
7:  qb_om.*
8:  qb_oz.*
9:  qb_pa11.*
10: qb_pa12.*
11: qb_fb.*
12: qb_pl1.*
13: qb_pl2.*
14: qb_pl3.*
15: qb_pl4.*
16: em_ba1.*
17: em_ba2.*
18: em_fm.*
19: em_hfq_bum.*
20: em_hfq_inz.*
21: em_hfq_rtd.*
22: em_hfq_bum1.*
23: em_hfq_bum2.*
24: em_hfq_bum3.*
25: em_hfq_bum4.*
26: em_hfq_bum5.*
27: em_hfq_bum6.*
28: em_hfq_bum7.*
29: em_hfq_bum8.*
30: em_hfq_inz1.*
31: em_hfq_inz2.*
32: em_hfq_rtd.*
33: em_hfq_rtd.*
34: em_bum1.*
35: em_bum2.*
36: em_bum3.*
37: em_bum4.*
38: em_bum5.*
39: em_bum6.*
40: em_bum7.*
41: em_fm.*
42: em_om.*
43: em_oz.*

```

| | | | | | | | |
|----|-------|------|--------|-----|----|------------|-------------|
| 12 | 508 | 50 | 7 | 755 | 9 | A0042536 | Sequence |
| 13 | 508 | 50.7 | 755 | 92 | 92 | HA223041 | |
| 14 | 442.6 | 44.2 | 800 | 86 | 86 | AF038567 | Homo sapi |
| 15 | 438.6 | 43.8 | 752 | 9 | 9 | AF037873 | Sequence |
| 16 | 438.6 | 43.8 | 752 | 9 | 9 | AX024591 | Sequence |
| 17 | 438.6 | 43.8 | 752 | 9 | 9 | AX024693 | Sequence |
| 18 | 438.6 | 43.8 | 752 | 93 | 93 | HS087459 | Human autoi |
| 19 | 437.6 | 43.7 | 755 | 9 | 9 | AF042537 | Sequence |
| 20 | 437.6 | 43.7 | 755 | 92 | 92 | HA0033149 | Homo sapi |
| 21 | 431 | 43.0 | 679 | 9 | 9 | AX024686 | Sequence |
| 22 | 431 | 43.0 | 679 | 92 | 92 | HA012833 | Homo sapi |
| 23 | 417 | 41.7 | 717 | 9 | 9 | AF042532 | Sequence |
| 24 | 119.2 | 11.9 | 21947 | 93 | 93 | HU061866FD | Homo sapien |
| 25 | 67.2 | 6.7 | 504 | 92 | 92 | HS1TR42 | Sequence |
| 26 | 65.4 | 6.5 | 16911 | 78 | 78 | AC087891 | Homo sapien |
| 27 | 60.6 | 6.0 | 38532 | 3 | 3 | SP094 | Streptococ |
| 28 | 58.8 | 5.9 | 62164 | 74 | 74 | AC098512 | Homo sapi |
| 29 | 56.6 | 5.6 | 60354 | 74 | 74 | AC069342 | Homo sapi |
| 30 | 56.6 | 5.6 | 149494 | 79 | 79 | AL182590 | Homo sapi |
| 31 | 56 | 5.6 | 166341 | 63 | 63 | AC013381 | Homo sapi |
| 32 | 55.8 | 5.6 | 80133 | 67 | 67 | AC021347 | Homo sapi |
| 33 | 55.2 | 5.5 | 73511 | 70 | 70 | AC025994 | Homo sapi |
| 34 | 55 | 5.5 | 126667 | 74 | 74 | AC015839 | Homo sapi |
| 35 | 54.4 | 5.4 | 244136 | 74 | 74 | AC069214 | Homo sapi |
| 36 | 53.8 | 5.4 | 12001 | 9 | 9 | AF048721 | Sequence |
| 37 | 53.8 | 5.4 | 13052 | 59 | 59 | HS1PI10P | Sequence |
| 38 | 53.8 | 5.4 | 47852 | 3 | 3 | MTV034 | Herpes simp |
| 39 | 53.4 | 5.4 | 72366 | 58 | 58 | AC022084 | Mycobacte |
| 40 | 53.8 | 5.4 | 152261 | 58 | 58 | HE10C | Homo sapi |
| 41 | 53.8 | 5.4 | 155963 | 69 | 69 | HF10C | Sequence |
| 42 | 53.2 | 5.3 | 135119 | 62 | 62 | AC011578 | Herpes simp |
| 43 | 53.2 | 5.3 | 202251 | 76 | 76 | AC074335 | Homo sapi |
| 44 | 52.8 | 5.3 | 62894 | 77 | 77 | AC084719 | Mus muscu |
| 45 | 52.8 | 5.3 | 188873 | 75 | 75 | AC073658 | Homo sapi |

ALIGNMENTS

| | |
|------------|---|
| RESULT | 1 |
| AR042535 | |
| Locus | AR042535 1002 bp DNA PAT |
| DEFINITION | Sequence 4 from patent US 5811519. |
| ACCESSION | AR042535 |
| VERSION | AR042535.1 GI:5963031 |
| KEYWORDS | . |
| SOURCE | .Unknown. |
| ORGANISM | unknown. unclassified, . (bases 1 to 1002) Lethe ,B., Lucas S., De Smet,C., Godelaine,D. and Boon-Falleur,T. LL-1 tumor specific genes Patent: US 5811519-A 4 22-SEP-1998; location,GenBankifiers 1..1002 organism= "unknown" |
| BASE COUNT | 179 a 302 c 336 g 185 t |

[illegible]


```

Db      612 CTTCATAGCTATCGCTCCCTGGCTTAGGAAATGCTGCACGACAGAGTGGCAGTTTCATTG 671
QY      913 taagagccatgattttatagctgaagaaacacagcttacaattttatctctataaaaa 972
Db      672 TGGGAACTGATGTTTCTGCTGCTGAGAGACAGACGCTTAACTGTTCTCTAGAAAA 731
QY      974 taagactgaactacaa 988
Db      742 TAAAGCTGAGCTACGA 747

RESULT 13
LOCUS   HSA223041 755 bp mRNA PRI 22-JUN 1998
DEFINITION Homo sapiens mRNA for LAGE-1a protein.
ACCESSION AJ223041
VERSION AJ223041.1 GI:1255994
KEYWORDS LAGE-1 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominiidae; Homo.
AUTHORS Lethe, B.G.
TITLE Direct Submission.
JOURNAL Submitted (06-JAN-1998) Lethe, B.G., Brussels Branch, Ludwig
Institute for Cancer Research, 74, avenue Hippocrate, B-1200
Bruxelles, BELGIUM
REFERENCE LAGE-1, a new gene with tumor specificity
AUTHORS Lotze, B., Lucas, S., Michaux, L., De Smet, C., Godolaine, D.,
Sottrano, A., De Plaen, E., and Boon, T.
LAGE-1, a new gene with tumor specificity
Int. J. Cancer 76 (b): 905-908 (1998)
94284662
FEATURES
location/Qualifiers
1..755
/organism="Homo sapiens"
/DB_xref="taxon:9606"
/chromosome="X"
/cell_line="IBX73-MFI"
/cell_type="Melanoma"
/map="q28"
/clone="IL-1"
/note="clone 4"
53..746
/gene="LAGE-1"
53..755
/gene="LAGE-1"
/note="related to NY-ESO-1"
/codon_start 1
/protein_id "CAA11044.1"
/db_xref="GI:1255995"
/translation="MVAEAGGIGGGTGDADGPGGGLPDGPGNAGGPKRGATGGRG
PRGAAAGASRRGGAPRRGGAGAAAGVWPPGAPPPSPPTQTHMPFSPDRA
ELVRLLSPDAAPLPPGAVLKDFVSCNLLFLRLTAADHRQLGLSSCLQQLSLIM
WTCGLPLVFLAQAAPSGRR"
730..745
/polyA_signal
746
/polyA_site
746
/gene="LAGE-1"
BASE COUNT 126 a 24 c 252 g 143 t
ORIGIN

Query Match 50.78; Score 508; DB 92; Length 755;
Best Local Similarity 76.58; Pred. No. 1; gc-83;
Matches 747; Conservative 0; Mismatches 0; Indels 229; Gaps 1;

QY 13 tctcatgagctgagctgctctctctgaagacagagagagctccagacacacagagag 72
Db 1 TCTCTGAGGCGCTGACTTCTCTCTCTGAGAGCCGGGAGAGGCTCCCGAGGCGCATGCAGGC 60
KEYWORDS
SOURCE human.

```

```

Db      61 CGAAGGCCAGGACACAGGCGTTCGACGCGCGATCTTCATGCGCCACAGAGGCGCTGGTAT 120
QY      133 tctcatgagctgagctgagctgagctgagctgagctgagctgagctgagctgagctgag 192
Db      121 TCTCTGATGAGGCGCAATGCTGAGGCGCGAGAGAGAGGCGGCTGCGCAGCGGCGGAG 180
QY      193 aatcccgagagagagagagagagagagagagagagagagagagagagagagagagag 252
Db      181 AGGTCCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY      253 tctcatgagctgagctgagctgagctgagctgagctgagctgagctgagctgagctgag 312
Db      241 TCGCATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 309
QY      313 gacagagagagagagagagagagagagagagagagagagagagagagagagagagagag 372
Db      301 GGACAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY      373 gctggtctcagagagagagagagagagagagagagagagagagagagagagagagagag 432
Db      361 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY      433 gaaagagagagagagagagagagagagagagagagagagagagagagagagagagagag 492
Db      421 GAAGGAGCTTACCGTGTCCGGCAACCTACTGTTAT 456
QY      493 aacgctcagagagagagagagagagagagagagagagagagagagagagagagagag 552
Db      457 ----- 456
QY      553 gaaagagagagagagagagagagagagagagagagagagagagagagagagagagagag 612
Db      457 ----- 456
QY      613 aagccagagagagagagagagagagagagagagagagagagagagagagagagagagag 672
Db      457 ----- 456
QY      673 gatggtctgagctgagctgagctgagctgagctgagctgagctgagctgagctgag 732
Db      457 -----CGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 491
QY      733 tctcatgagctgagctgagctgagctgagctgagctgagctgagctgagctgagctgag 792
Db      492 TCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 551
QY      793 tgcagagagagagagagagagagagagagagagagagagagagagagagagagagagag 852
Db      552 TGGGCGTGTGTTTGGCTAGGCTCTCTGAGGCTAGAGGCTAGAGGCTAGAGGCTAGAGGCT 611
QY      853 ctctcatgagctgagctgagctgagctgagctgagctgagctgagctgagctgagctgag 912
Db      612 CTTCCTAGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 671
QY      913 tgaagagagagagagagagagagagagagagagagagagagagagagagagagagagag 972
Db      672 TGGGCGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 731
QY      973 taagactgaactacaa 988
Db      732 TAAAGCTGAGCTACGA 747

```

```

RESULT 14
LOCUS   AF038567 806 bp mRNA PRI 16-JAN-1999
DEFINITION Homo sapiens cancer antigen-3 and cancer antigen-3-related mRNA,
complete cds.
ACCESSION AF038567
VERSION AF038567.1 GI:4104676
KEYWORDS
SOURCE human.

```


Genware version 4.5
Copyright (c) 1993 - 2000 CompuLink Ltd.

0M nucleotide nucleotide search, using sw model

Run on: April 10, 2001, 23:54:02 : Search time 4786.75 Seconds
(without alignments)
2311631 Million cell updates/sec

Initial: US-09-341-829A-4
Post-test score: 1002
Sequence: 1 test query extracted (q.....ctacgattccgagagaaaa 1002

Search method: HENKLEY NUC
Gapop 10.0 : Gapext 1.0

Searches: 962537 seqs, 436804070 residues

Total number of hits satisfying chosen parameters: 19247034

Minimum hit seq length: 0

Maximum hit seq length: 2000000000

Post processing: Maximum Match 9%

Maximum Match 100%

Listing first 45 summaries

| Database : | EST.* |
|----------------|-------|
| 1: qb_est1.* | |
| 2: qb_est2.* | |
| 3: qb_est3.* | |
| 4: qb_est4.* | |
| 5: qb_est5.* | |
| 6: qb_est6.* | |
| 7: qb_est7.* | |
| 8: qb_est8.* | |
| 9: qb_est9.* | |
| 10: qb_est10.* | |
| 11: qb_est11.* | |
| 12: qb_est12.* | |
| 13: qb_est13.* | |
| 14: qb_est14.* | |
| 15: qb_est15.* | |
| 16: qb_est16.* | |
| 17: qb_est17.* | |
| 18: qb_est18.* | |
| 19: qb_est19.* | |
| 20: qb_est20.* | |
| 21: qb_est21.* | |
| 22: qb_est22.* | |
| 23: qb_est23.* | |
| 24: qb_est24.* | |
| 25: qb_est25.* | |
| 26: qb_est26.* | |
| 27: qb_est27.* | |
| 28: qb_est28.* | |
| 29: qb_est29.* | |
| 30: qb_est30.* | |
| 31: qb_est31.* | |
| 32: qb_est32.* | |
| 33: qb_est33.* | |
| 34: qb_est34.* | |
| 35: qb_est35.* | |
| 36: qb_est36.* | |
| 37: qb_est37.* | |
| 38: qb_est38.* | |
| 39: qb_est39.* | |
| 40: qb_est40.* | |
| 41: qb_est41.* | |
| 42: qb_est42.* | |
| 43: qb_est43.* | |
| 44: qb_est44.* | |
| 45: qb_est45.* | |
| 46: qb_est46.* | |
| 47: qb_est47.* | |

| | |
|------------------|---------------|
| 44: qb_est44.* | em_esthum10.* |
| 45: qb_est45.* | em_esthum11.* |
| 46: qb_est46.* | em_esthum12.* |
| 47: qb_est47.* | em_esthum13.* |
| 48: qb_est48.* | em_esthum14.* |
| 49: qb_est49.* | em_esthum15.* |
| 50: qb_est50.* | em_esthum16.* |
| 51: qb_est51.* | em_esthum17.* |
| 52: qb_est52.* | em_esthum18.* |
| 53: qb_est53.* | em_esthum19.* |
| 54: qb_est54.* | em_esthum20.* |
| 55: qb_est55.* | em_esthum21.* |
| 56: qb_est56.* | em_esthum22.* |
| 57: qb_est57.* | em_esthum23.* |
| 58: qb_est58.* | em_esthum24.* |
| 59: qb_est59.* | em_esthum25.* |
| 60: qb_est60.* | em_esthum26.* |
| 61: qb_est61.* | em_esthum27.* |
| 62: qb_est62.* | em_esthum28.* |
| 63: qb_est63.* | em_esthum29.* |
| 64: qb_est64.* | em_esthum30.* |
| 65: qb_est65.* | em_esthum31.* |
| 66: qb_est66.* | em_esthum32.* |
| 67: qb_est67.* | em_esthum33.* |
| 68: qb_est68.* | em_esthum34.* |
| 69: qb_est69.* | em_esthum35.* |
| 70: qb_est70.* | em_esthum36.* |
| 71: qb_est71.* | em_esthum37.* |
| 72: qb_est72.* | em_esthum38.* |
| 73: qb_est73.* | em_esthum39.* |
| 74: qb_est74.* | em_esthum40.* |
| 75: qb_est75.* | em_esthum41.* |
| 76: qb_est76.* | em_esthum42.* |
| 77: qb_est77.* | em_esthum43.* |
| 78: qb_est78.* | em_esthum44.* |
| 79: qb_est79.* | em_esthum45.* |
| 80: qb_est80.* | em_esthum46.* |
| 81: qb_est81.* | em_esthum47.* |
| 82: qb_est82.* | em_esthum48.* |
| 83: qb_est83.* | em_esthum49.* |
| 84: qb_est84.* | em_esthum50.* |
| 85: qb_est85.* | em_esthum51.* |
| 86: qb_est86.* | em_esthum52.* |
| 87: qb_est87.* | em_esthum53.* |
| 88: qb_est88.* | em_esthum54.* |
| 89: qb_est89.* | em_esthum55.* |
| 90: qb_est90.* | em_esthum56.* |
| 91: qb_est91.* | em_esthum57.* |
| 92: qb_est92.* | em_esthum58.* |
| 93: qb_est93.* | em_esthum59.* |
| 94: qb_est94.* | em_esthum60.* |
| 95: qb_est95.* | em_esthum61.* |
| 96: qb_est96.* | em_esthum62.* |
| 97: qb_est97.* | em_esthum63.* |
| 98: qb_est98.* | em_esthum64.* |
| 99: qb_est99.* | em_esthum65.* |
| 100: qb_est100.* | em_esthum66.* |
| 101: qb_est101.* | em_esthum67.* |
| 102: qb_est102.* | em_esthum68.* |
| 103: qb_est103.* | em_esthum69.* |
| 104: qb_est104.* | em_esthum70.* |
| 105: qb_est105.* | em_esthum71.* |
| 106: qb_est106.* | em_esthum72.* |
| 107: qb_est107.* | em_esthum73.* |
| 108: qb_est108.* | em_esthum74.* |
| 109: qb_est109.* | em_esthum75.* |
| 110: qb_est110.* | em_esthum76.* |
| 111: qb_est111.* | em_esthum77.* |
| 112: qb_est112.* | em_esthum78.* |
| 113: qb_est113.* | em_esthum79.* |
| 114: qb_est114.* | em_esthum80.* |
| 115: qb_est115.* | em_esthum81.* |
| 116: qb_est116.* | em_esthum82.* |

```

117: qb_est48:*
118: qb_est49:*
119: qb_est50:*
120: qb_est51:*
121: qb_est52:*
122: qb_est53:*
123: qb_est54:*
124: qb_est55:*
125: qb_est56:*
126: qb_est57:*
127: qb_est58:*
128: qb_est59:*
129: qb_est60:*
130: qb_est61:*
131: qb_est62:*
132: qb_est63:*
133: qb_est64:*
134: qb_est65:*
135: qb_est66:*
136: qb_est67:*
137: qb_est68:*
138: qb_est69:*
139: qb_est70:*
140: qb_est71:*
141: qb_est72:*
142: qb_est73:*
143: qb_est74:*
144: qb_est75:*
145: qb_est76:*
146: qb_est77:*
147: qb_est78:*
148: qb_est79:*
149: qb_est80:*
150: qb_est81:*
151: qb_est82:*
152: qb_est83:*
153: qb_est84:*
154: qb_est85:*
155: qb_est86:*
156: qb_est87:*
157: qb_est88:*
158: qb_est89:*
159: qb_est90:*
160: qb_est91:*
161: qb_est92:*
162: qb_est93:*
163: qb_est94:*
164: qb_est95:*
165: qb_est96:*
166: qb_est97:*
167: qb_est98:*
168: qb_est99:*
169: qb_est100:*
170: qb_est101:*
171: qb_est102:*
172: qb_est103:*
173: qb_est104:*
174: qb_est105:*
175: qb_est106:*
176: qb_est107:*
177: qb_est108:*
178: qb_est109:*
179: qb_est110:*
180: qb_est111:*
181: qb_est112:*
182: qb_est113:*
183: qb_est114:*
184: qb_est115:*
185: qb_est116:*
186: qb_est117:*
187: qb_est118:*
188: qb_est119:*
189: qb_est120:*

```

```

190: em_oss_pln1:*
191: em_oss_pln2:*
192: em_oss_pro:*
193: em_oss_rod1:*
194: em_oss_rod2:*
195: em_oss_rod3:*
196: em_oss_rod4:*
197: em_oss_rod5:*
198: em_oss_vrt1:*
199: em_oss_vrt2:*
200: em_oss_vrt3:*
201: qb_oss1:*
202: qb_oss2:*
203: qb_oss3:*
204: qb_oss4:*
205: qb_oss5:*
206: qb_oss6:*
207: qb_oss7:*
208: qb_oss8:*
209: qb_oss9:*
210: qb_oss10:*
211: qb_oss11:*
212: qb_oss12:*
213: qb_oss13:*
214: qb_oss14:*
215: qb_oss15:*
216: qb_oss16:*
217: qb_oss17:*
218: qb_oss18:*
219: qb_oss19:*
220: qb_oss20:*
221: qb_oss21:*
222: qb_oss22:*
223: qb_oss23:*
224: qb_oss24:*
225: qb_oss25:*
226: qb_oss26:*
227: qb_oss27:*
228: qb_oss28:*
229: qb_oss29:*
230: qb_oss30:*
231: qb_oss31:*
232: qb_oss32:*
233: qb_oss33:*
234: qb_oss34:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | Query Match % | Length | DB | ID | Description |
|--------|-----|-------|---------------|--------|-----|----------|--------------------|
| | 1 | 579.2 | 57.8 | 629 | 165 | HE274123 | HE274123 601385146 |
| | 2 | 528.6 | 52.8 | 586 | 166 | HE390259 | HE390259 601385141 |
| | 3 | 467.8 | 46.7 | 739 | 166 | HE291232 | HE291232 601385820 |
| | 4 | 461.8 | 46.1 | 559 | 165 | HE276104 | HE276104 601144452 |
| | 5 | 393.2 | 39.2 | 800 | 166 | HE408892 | HE408892 601303769 |
| | 6 | 373.6 | 37.3 | 674 | 166 | HE388562 | HE388562 601281855 |
| | 7 | 373.2 | 37.2 | 646 | 165 | HE410697 | HE410697 601309763 |
| | 8 | 357.8 | 35.7 | 511 | 166 | HE387918 | HE387918 601282166 |
| | 9 | 346.6 | 34.6 | 383 | 166 | HE410952 | HE410952 601303669 |
| | 10 | 289.2 | 28.9 | 359 | 17 | A1218222 | A1218222 601303232 |
| | 11 | 269.8 | 26.9 | 281 | 10 | AA634317 | AA634317 257860518 |
| | 12 | 227.8 | 22.7 | 639 | 166 | HE387902 | HE387902 601271781 |
| | 13 | 227.6 | 22.7 | 952 | 166 | HE385880 | HE385880 601275448 |
| | 14 | 86.6 | 8.6 | 321 | 115 | AW436516 | AW436516 76782 MAR |
| | 15 | 85.6 | 8.5 | 148 | 13 | AA866133 | AA866133 0154309.S |
| | 16 | 75.8 | 7.5 | 426 | 10 | AA682581 | AA682581 2120003.S |
| | 17 | 75.6 | 7.5 | 447 | 154 | HE4542 | HE4542 2701008.S1 |
| | 18 | 73.2 | 7.3 | 693 | 138 | HE27845 | HE27845 601364266 |

RESULT 3

| | | | | | |
|------------|----------------|------------|--------------|------------|-----------------|
| BE391232 | BE391242 | 739 bp | mRNA | EST | 21-Jul-2000 |
| LOCUS | CG1285923F1 | NTL_MGC_44 | Homo sapiens | cdna clone | IMAGE:4607708-5 |
| DEFINITION | mRNA sequence. | | | | |
| ACCESSION | BE391232 | | | | |
| VERSION | BE391242.1 | GI:5346597 | | | |
| KEYWORDS | EST. | | | | |

| | |
|-----------|---|
| ORGANISM | Homo sapiens |
| REFERENCE | Eukaryota; Metazoa; Chordata; Crustacea; Mammalia; Eutheria; Primates; Catarrhini (bases 1 to 749) |
| ARTICLE | g1c-25c-h1g-1-3-4-5-6-h1g-1-3-4-5-6 |
| TITLE | National Institutes of Health, Marburg |
| JOURNAL | Unpublished (1996) |
| COMMENT | Contact: Robert Strausberg, ph.D. Tel.: (301) 496-1950 Email: Robert_Strausberg@nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Ling He cDNA Library Arrayed by: The I.M.C. DNA Sequencing by: Incyte Genomics Clone Distribution: MGC clone distributed through the I.M.A.G.E. Consortium http://image.llnl.gov Plate: LHC259 row: a column: 09 High quality sequence stop: 631. |

[illegible]

DB 142 TTTCGGATTAAATTTTGAAGACCCGCCCAAAAGGGATTTTCCAAAAGCTCTTTTTC 74
QY 754 agcgaattttccctatgatatgatacagcagtgactttctacccgtcttttttagctcaag 814
DB 72 TT
DB 72 AGGAGCTTTCCCTGTGTGATGTGGATCCGCCAGTCTTTTCCCGCTTTTTRGGCTCAGC 13
QY 814 ctccctcaaggc 825
DB 12 GCGCGCGAGGG 1

Search completed: May 1, 2001, 04:03:56
Job time: 11.494 sec

$$S^{-1} \psi = \psi \quad \text{for } \psi \in \mathcal{H}_0 \quad \text{and} \quad S^{-1} \psi = -\psi \quad \text{for } \psi \in \mathcal{H}_1.$$

5. Theorem of the Rank

[illegible]

May 1, 2004, 01:52:17 : Search time 158.99 seconds
(without alignments)
1158,699 Million cell updates/sec

111100: US 09 041 829A-4
 Perfect score: 100%
 Summary: 1 perfect, 0 above, 0 below
 of good predictions: 100%

IDENTITY NP's
Gapref 10.0 , Gapref 1.0

searched: 86262 seqs, 8750344 residues
total number of hits satisfying chosen parameters: 605242

$$\begin{array}{l} \text{M}[\text{triminim}]\text{ls} \text{ seq} \vdash \text{seq} \vdash \text{ls}; \quad 0 \\ \text{M}[\text{ximinim}]\text{ls} \text{ seq} \vdash \text{seq} \vdash \text{ls}; \quad 2 \text{ seq seq seq seq} \end{array}$$

Maximum Match 100%

Listed first 45 summaries

[illegible]

4. $\mathcal{H}^1(\mathbb{R}^n) \subset \mathcal{H}^1(\mathbb{R}^n)$ (see [1, Ch. 1, § 1.1]).
5. $\mathcal{H}^1(\mathbb{R}^n) \subset \mathcal{H}^1(\mathbb{R}^n)$ (see [1, Ch. 1, § 1.1]).
6. $\mathcal{H}^1(\mathbb{R}^n) \subset \mathcal{H}^1(\mathbb{R}^n)$ (see [1, Ch. 1, § 1.1]).

Proof. N_{α} is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

[illegible]

| | |
|--------------------|--|
| Sequence 7, Appl | |
| Sequence 1, Appl | |
| Sequence 8, Appl | |
| Sequence 9, Appl | |
| Sequence 10, Appl | |
| Sequence 11, Appl | |
| Sequence 12, Appl | |
| Sequence 13, Appl | |
| Sequence 14, Appl | |
| Sequence 15, Appl | |
| Sequence 16, Appl | |
| Sequence 17, Appl | |
| Sequence 18, Appl | |
| Sequence 19, Appl | |
| Sequence 20, Appl | |
| Sequence 21, Appl | |
| Sequence 22, Appl | |
| Sequence 23, Appl | |
| Sequence 24, Appl | |
| Sequence 25, Appl | |
| Sequence 26, Appl | |
| Sequence 27, Appl | |
| Sequence 28, Appl | |
| Sequence 29, Appl | |
| Sequence 30, Appl | |
| Sequence 31, Appl | |
| Sequence 32, Appl | |
| Sequence 33, Appl | |
| Sequence 34, Appl | |
| Sequence 35, Appl | |
| Sequence 36, Appl | |
| Sequence 37, Appl | |
| Sequence 38, Appl | |
| Sequence 39, Appl | |
| Sequence 40, Appl | |
| Sequence 41, Appl | |
| Sequence 42, Appl | |
| Sequence 43, Appl | |
| Sequence 44, Appl | |
| Sequence 45, Appl | |
| Sequence 46, Appl | |
| Sequence 47, Appl | |
| Sequence 48, Appl | |
| Sequence 49, Appl | |
| Sequence 50, Appl | |
| Sequence 51, Appl | |
| Sequence 52, Appl | |
| Sequence 53, Appl | |
| Sequence 54, Appl | |
| Sequence 55, Appl | |
| Sequence 56, Appl | |
| Sequence 57, Appl | |
| Sequence 58, Appl | |
| Sequence 59, Appl | |
| Sequence 60, Appl | |
| Sequence 61, Appl | |
| Sequence 62, Appl | |
| Sequence 63, Appl | |
| Sequence 64, Appl | |
| Sequence 65, Appl | |
| Sequence 66, Appl | |
| Sequence 67, Appl | |
| Sequence 68, Appl | |
| Sequence 69, Appl | |
| Sequence 70, Appl | |
| Sequence 71, Appl | |
| Sequence 72, Appl | |
| Sequence 73, Appl | |
| Sequence 74, Appl | |
| Sequence 75, Appl | |
| Sequence 76, Appl | |
| Sequence 77, Appl | |
| Sequence 78, Appl | |
| Sequence 79, Appl | |
| Sequence 80, Appl | |
| Sequence 81, Appl | |
| Sequence 82, Appl | |
| Sequence 83, Appl | |
| Sequence 84, Appl | |
| Sequence 85, Appl | |
| Sequence 86, Appl | |
| Sequence 87, Appl | |
| Sequence 88, Appl | |
| Sequence 89, Appl | |
| Sequence 90, Appl | |
| Sequence 91, Appl | |
| Sequence 92, Appl | |
| Sequence 93, Appl | |
| Sequence 94, Appl | |
| Sequence 95, Appl | |
| Sequence 96, Appl | |
| Sequence 97, Appl | |
| Sequence 98, Appl | |
| Sequence 99, Appl | |
| Sequence 100, Appl | |

| | | | | | | | |
|---|---|------|---------|---|----------------|----|-------------------|
| c | 7 | 5, 4 | 1, 2001 | 1 | 98-08-458-568A | 11 | Sequence 11, Appl |
| b | 8 | 4, 3 | 2580 | 4 | 98-09-050-803 | 2 | Sequence 2, Appl |
| e | 9 | 4, 3 | 5452 | 2 | 98-09-100-114 | 1 | Sequence 1, Appl |

| | | | | | | | |
|----|----|------|-------|---|----------------|----|-------------------|
| 11 | 48 | 4, 8 | 10596 | 1 | 08-07-885-971 | 15 | Sequence 15, Appl |
| 12 | 48 | 4, 8 | 10596 | 1 | 08-08-037-79A | 15 | Sequence 15, Appl |
| 13 | 48 | 4, 8 | 10596 | 1 | 08-08-194-088B | 15 | Sequence 15, Appl |

[illegible]

| | | | | | | |
|----|------|-----|-------|---|-------------------|-------------------|
| 18 | 45.2 | 4.5 | 4.24 | 4 | US-09-128-537-7 | Sequence 7, Appl |
| 19 | 45.2 | 4.5 | 4.233 | 4 | US-09-128-665-16 | Sequence 16, Appl |
| 20 | 44 | 4.4 | 4.30 | 4 | US-08-758-662-4 | Sequence 4, Appl |
| 21 | 43.9 | 4.4 | 4.50 | 2 | US-08-609-4420-17 | Sequence 17, Appl |

| | | | | | | | |
|---|-----|------|-----|------|---|-------------------|-------------------|
| c | 2.3 | 43.4 | 4.3 | 801 | 2 | US-08-770-579-16 | Sequence 16, Appl |
| e | 2.3 | 43.4 | 4.3 | 801 | 4 | US-08-757-669A-16 | Sequence 16, Appl |
| d | 2.4 | 43.2 | 4.3 | 2140 | 1 | US-08-334-698-1 | Sequence 1, Appl |

[illegible]

| | | | | | | | |
|----|-------|-----|-------|---|----|-----------------|--------------------|
| 28 | 4.5-2 | 4.3 | 2.140 | 2 | US | 08-722-1-00-1 | Sequence 1, App.1 |
| 29 | 4.3-2 | 4.3 | 2.140 | 3 | US | 08-722-1-00-1 | Sequence 1, App.1 |
| 30 | 4.3-2 | 4.3 | 2.140 | 3 | US | 09-206-6-99-1 | Sequence 1, App.1 |
| 31 | 4.3-2 | 4.3 | 2.140 | 5 | US | 01-089, 04-04-1 | Sequence 1, App.1 |
| 42 | 4.2-6 | 4.3 | 6.453 | 1 | US | 08-306-6-01B-14 | Sequence 14, App.1 |
| 43 | 4.2-6 | 4.3 | 6.453 | 4 | US | 09-209-6-00-10 | Sequence 10, App.1 |
| 44 | 4.2-6 | 4.3 | 6.453 | 4 | US | 09-356-6-02-8 | Sequence 2, App.1 |
| 45 | 4.2-1 | 4.2 | 5.622 | 3 | US | 08-678-1-00-2 | Sequence 2, App.1 |
| 46 | 4.2-4 | 4.2 | 5.677 | 3 | US | 08-678-1-06-1 | Sequence 14, App.1 |
| 47 | 4.1-8 | 4.2 | 7.218 | 1 | US | 08-342-2-13-14 | Sequence 1, App.1 |
| 48 | 4.1-6 | 4.2 | 1.269 | 4 | US | 08-306-2-18-1 | Sequence 1, App.1 |
| 49 | 4.1-6 | 4.2 | 1.276 | 1 | US | 08-740-1-16-1 | Sequence 1, App.1 |
| 40 | 4.1-6 | 4.2 | 1.776 | 1 | US | 08-722-0-01-29 | Sequence 29, App.1 |
| 41 | 4.1-4 | 4.1 | 1.608 | 4 | US | 09-222-0-34B-19 | Sequence 7, App.1 |
| 42 | 4.1-2 | 4.1 | 4.20 | 4 | US | 09-165-2-04-7 | Sequence 8, App.1 |
| 43 | 4.1 | 4.1 | 3.19 | 4 | US | 09-165-2-04-8 | Sequence 1, App.1 |
| 44 | 4.1 | 4.1 | 3.53 | 2 | US | 08-745-8-75-1 | Sequence 1, App.1 |
| 45 | 4.1 | 4.1 | 3.53 | 3 | US | 08-745-8-75-1 | Sequence 1, App.1 |

APPENDIX

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```



```

1 GENERAL INFORMATION:
2 APPLICANT: GADSWILL, Paul J 10000, Nathaniel A. Mark, Melanie R.
3 TITLE OF INVENTION: SINGLE-CHAIN HETEROZYCLE GROWTH FACTOR VARIANTS
4 NUMBER OF SEQUENCES: 21
5 CORRESPONDENCE ADDRESS:
6 APPLICANT: GADSWILL, Paul J 10000, Nathaniel A. Mark, Melanie R.
7 ADDRESS: Genentech, Inc.
8 STREET: 460 Point San Bruno Blvd
9 CITY: South San Francisco
10 STATE: California
11 COUNTRY: USA
12 ZIP: 94080
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: 5.25 Inch, 400 KB floppy disk
15 COMPILER: IBM PC compatible
16 OPERATING SYSTEM: PC-DOS/MS-DOS
17 SOFTWARE: PATTB (Genentech)
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/08/194 0000
20 FILING DATE: 09 FEB 1994
21 CLASSIFICATION: 530
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: 07/984811
24 FILING DATE: 18 MAY 1992
25 ALTERNATIVE INFORMATION:
26 NAME: Gallages, R. Thomas
27 RESEQUENCING NUMBER: 52,652
28 REFERENCE SEQUENCE NUMBER: 75511
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: 415/225 2614
31 TELEFAX: 415/952 0881
32 FAX: 916/471 7168
33 INFORMATION FOR SEQ ID NO: 15:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 1056 bases
36 TYPE: nucleic acid
37 STRAIN/GENESS: Str8le
38 TISSUE: Throat
39 SEQ ID NO: 15
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
```

[illegible]

```

STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: pc DOS/MS DOS
SOFTWARE: patlin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 08/08/194-087
FILING DATE: 18 MAY-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dregor, Gindot R.
REGISTRATION NUMBER: 43,055
REFERENCE/DOCKET NUMBER: 779
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-4216
TELEFAX: 415/952-9881
TELEX: 910/471-7168
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 10596 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-194-087-15

Query Match          4.8%      Score 48;   BB 2;    Length 10596;
Best Local Similarity 51.9%;   P-val. No. 0.028;
Matches 108;   Conserved 0;   Mismatches 100;   Indels 0;   Gaps 0;

QY 46 qqacaaagqctccgaadgcatacaagcagaatcccaggacacgagggatctcgagggcga 105
   || ||| || |||| | || | || | || | || | || | || | || | || | || |
DB 2734 GGAGGAGGGGATGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGG 2793

QY 106 tctactatgaccaggggtccctgaccttcatgatgccctggaggttcctgagggc 165
   || | || | || || | || || | || | || | || | || | || | || | || |
DB 2794 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2853

QY 166 atgaagagagagatgcacacagagcagcaggaattccctggagcagagagagcctg 225
   || || | || | || | || | || | || | || | || | || | || | || | || |
DB 2854 AGAGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAGAGGAG 2913

QY 226 qqqqccgaagagagagagagcctccgcagagagat 253
   || | || | || | || | || | || | || | || | || | || | || | || |
DB 2914 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2941

RESULT 15
PCT-US93-04548-15
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.; Gubowskii, Paul J., Lasker, Nathan A., Mark, Michael
TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSER: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: pc DOS/MS DOS
SOFTWARE: patlin (Genentech)
CURRENT APPLICATION DATA:

```


XX New isolated LAGE-1 tumour associated nucleic acids - used to
 PT develop products for the diagnosis and treatment of LAGE-1
 PI associated disorders, particularly tumours
 XX
 PS Claim 1: Page 50-52; 7pp; English.
 XX
 CC The present sequence encodes LAGE-1 tumour associated protein (TAP).
 CC The present invention also describes: (1) a method for treating a
 CC subject with a disorder characterised by expression of a LAGE-1 nucleic
 CC acid molecule or an expression product, comprising administering to the
 CC subject autologous cytolytic T cells to ameliorate the disorder, where
 CC the cytolytic T cells are specific for complexes of an HLA molecule and
 CC a LAGE-1 TAP or an immunogenic fragment; (2) a method for treating a
 CC subject with a disorder characterised by expression of a LAGE-1 nucleic
 CC acid molecule or an expression product, comprising administering a
 CC LAGE-1 TAP or an immunogenic fragment to ameliorate the disorder; and
 CC (3) a method for selectively enriching a population of T cells with
 CC cytolytic T cells specific for a LAGE-1 TAP comprising contacting an
 CC isolated population of T cells with an agent presenting a complex of a
 CC LAGE TAP or an immunogenic fragment and a HLA presenting molecule to
 CC selectively enrich the isolated population of T cells with the cytolytic
 CC T cells. The methods and products from the present invention can be used
 CC for the diagnosis and treatment of LAGE-1 associated disorders,
 CC particularly tumours.
 XX
 SQ Sequence 1002 BP; 179 A; 302 C; 336 G; 185 T; 0 other;

Query Match 100.0%; Score 14; DB 19; Length 1002;
 Best Local Similarity 100.0%; pred. No. 1.5e+02;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tetactccatcgcac 14
 DB 1 tetactccatcgcac 14

RESULT 2
 X12007/c
 ID X12007 standard: DNA; 133 BP.
 AC
 AT
 XX X12007;
 DF 40-MAR-1999 (first entry)
 XX
 DE Human biallelic polymorphic DNA fragment W1-18427.
 KW Polymorphism; biallelic; human; forensics; paternity testing; disease;
 KW detection; phenotypic typing; characteristic; infection; hereditary;
 KW autoimmune disease; cancer; inflammation; drug; therapy; medication;
 KW treatment; marker; ss.
 XX
 OS Homo sapiens.
 XX
 PN W9820145-A2.
 XX
 PD 14-MAY-1998.
 XX
 PF 05-NOV-1997; 97W0-0S20313.
 XX
 PR 06-NOV-1996; 96US-0030455.
 XX
 PA (WHEED) WHITEHEAD INST BIOMEDICAL RES.
 XX
 PI Hudson L, Lander ES, Wang D;
 XX
 PR WPI; 1998-286974/25.
 XX
 PT New isolated nucleic acid segments from the human genome - used for
 PT determining polymorphic forms for use in e.g. forensics, paternity
 PT testing or phenotypic typing for disease
 XX

PS Claim 1: Page 212; 410pp; English.
 XX
 CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
 CC markers which have been isolated using the primers represented in
 CC X09121-X10268. The base sequence of the polymorphic site is indicated by
 CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
 CC methods for determining polymorphic forms in an individual for use in
 CC e.g. forensics, paternity testing or for phenotypic typing for diseases
 CC such as amniotubularia, diabetes insipidus, Lesch-Nyhan syndrome,
 CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
 CC hypercholesterolemia, polycystic kidney disease, hereditary
 CC spherocytosis, von Willebrand's disease, tuberculous sclerosis, hereditary
 CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers Danlos
 CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
 CC autoimmune diseases, inflammation, cancer, diseases of the nervous
 CC system, infection by pathogenic microorganisms, and characteristics such
 CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
 CC endurance, fertility, and susceptibility or reactivity to particular
 CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
 CC segments can also be used to produce medicaments for the treatment or
 CC prophylaxis of such diseases.
 XX
 SQ Sequence 133 BP; 40 A; 27 C; 35 G; 30 T; 1 other;

Query Match 92.9%; Score 13; DB 19; Length 133;
 Best Local Similarity 100.0%; pred. No. 4.8e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tetactccatcgcac 13
 DB 131 TCTGCTTCGCAT 119

RESULT 3
 C53862/c
 ID C53862 standard: DNA; 501 BP.
 AC
 AT
 XX C53862;
 DF 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID No: 75924.
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 24-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 24-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.

PR 21-OCT-1999; 990S-0160768.
 PR 21-OCT-1999; 990S-0160770.
 PR 21-OCT-1999; 990S-0160814.
 PR 21-OCT-1999; 990S-0160815.
 PR 22-OCT-1999; 990S-0160980.
 PR 22-OCT-1999; 990S-0160981.
 PR 22-OCT-1999; 990S-0160989.
 PR 25-OCT-1999; 990S-0161404.
 PR 25-OCT-1999; 990S-0161405.
 PR 25-OCT-1999; 990S-0161406.
 PR 26-OCT-1999; 990S-0161359.
 PR 26-OCT-1999; 990S-0161360.
 PR 26-OCT-1999; 990S-0161361.
 PR 28-OCT-1999; 990S-0161920.
 PR 28-OCT-1999; 990S-0161952.
 PR 28-OCT-1999; 990S-0161953.
 PR 29-OCT-1999; 990S-0162142.

Query Match 92.9%; Score 13; DB 21; Length 501;
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ctgcctccgcac 14
 |||||

DB 254 CTGCTCCGCAC 242

RESULT 4
 V89990; c
 ID V89990 standard; cDNA; 624 BP.

AC V89990;

XX 15-FEB-1999 (first entry)

DE EST clone CW707.

XX Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
 KW tissue growth protein; inhibiting chemotaxis; chemokinesis; haemostatic;
 KW cooperator; fibrin; thrombolytic; anti-inflammatory; cadherin; anti tumour;
 KW gene therapy; ss.

OS Homo sapiens.

XX W99845436-A2.

XX 15-OCT-1998.

XX 10-APR-1998; 98W0-0506955.

XX 10-APR-1997; 970S-0838821.

XX (GENY) GENETICS INST INC.

XX Anagnostou M, Jacobs Z, Lavallie E, McCoy JM, Murburg D,

XX Racine LA, Spaulding V, Treacy M;

XX WPI; 1999-070077/06

XX New polynucleotides encoding human secreted proteins - derived from
 PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
 PT ovary, pituitary, retina and colon cDNA libraries.

XX Claim 1: Page 396; 618pp; English.

XX The present sequence represents a human expressed sequence tag (EST).
 CC the polynucleotide, which is a secreted EST, and the encoded protein
 CC are predicted to have useful biological activities which would make
 CC them suitable for treating, preventing or ameliorating medical
 CC conditions in humans and animals, although no supporting data is
 CC given. Suggested activities include nutritional activity, immune
 CC stimulating or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activity/inhibin activity.

CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The polynucleotide may also be useful for gene therapy.
 XX
 SQ Sequence 624 BP; 175 A; 158 C; 129 G; 162 T; 0 other;

Query Match 92.9%; Score 13; DB 29; Length 624;
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ctgcctccgcac 14
 |||||

DB 604 CTGCTCCGCAC 592

RESULT 5

ID N90091 standard; cDNA; 885 BP.

AC N90091;

XX 01-NOV-1989 (first entry)

DE Gene 519 cDNA.

KW cDNA; gene 519; T-cell peptide; functional, activated
 KW cells; human; detecting tumours.

XX Homo sapiens.

PH Key location/Qualifiers

PT CDS 279..666

ET /*tag a

XX EP420806-A.

XX 21-JUN-1989.

XX 08-DEC-1988; 88EP-0120556.

XX 15-DEC-1987; 87US-0132926.

XX (STRO) LELAND STANFORD JR UNIV.

XX Krensky AM, Davis M, Schall T, Jonstra J;

XX WPI; 1989-179693/25.

XX P-PSDB; P90346.

XX New cDNA for gene 519 encoding new T-cell peptide
 PT only in functional, activated cells, useful for
 PT detecting tumours and assessing T-cell status.

XX Claim: page 8; 12pp; English.

XX cDNA (see vntesp. P90346) has the sequence reported for gene 519.
 CC It encodes a peptide associated with normal T-cells, and can detect
 CC functional, activated T cells to distinguish between neoplastic cells.

XX Sequence 885 BP; 201 A; 260 C; 240 G; 184 T; 0 other;

Query Match 92.9%; Score 13; DB 10; Length 885;
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ctgcctccgcac 14
 |||||

DB 49 ctgcctccgcac 61


```

PR 20-AUG-1999; 990S-0149929;
PR 23-AUG-1999; 990S-0149902;
PR 23-AUG-1999; 990S-0149930;
PR 25-AUG-1999; 990S-0150566;
PR 26-AUG-1999; 990S-0150894;
PR 27-AUG-1999; 990S-0151065;
PR 27-AUG-1999; 990S-0151066;
PR 27-AUG-1999; 990S-0151090;
PR 30-AUG-1999; 990S-0151303;
PR 31-AUG-1999; 990S-0151438;
PR 01-SEP-1999; 990S-0151970;
PR 07-SEP-1999; 990S-0152363;
PR 10-SEP-1999; 990S-0152070;
PR 13-SEP-1999; 990S-0153758;
PR 15-SEP-1999; 990S-0154018;
PR 15-SEP-1999; 990S-0154039;
PR 22-SEP-1999; 990S-0154779;
PR 23-SEP-1999; 990S-0155139;
PR 23-SEP-1999; 990S-0155486;
PR 24-SEP-1999; 990S-0155659;
PR 28-SEP-1999; 990S-0156458;
PR 29-SEP-1999; 990S-0156596;
PR 04-OCT-1999; 990S-0157117;
PR 05-OCT-1999; 990S-0157753;
PR 06-OCT-1999; 990S-0157865;
PR 07-OCT-1999; 990S-0158029;
PR 08-OCT-1999; 990S-0158272;
PR 12-OCT-1999; 990S-0158469;
PR 13-OCT-1999; 990S-0159293;
PR 13-OCT-1999; 990S-0159294;
PR 13-OCT-1999; 990S-0159295;
PR 14-OCT-1999; 990S-0159429;
PR 14-OCT-1999; 990S-0159430;
PR 14-OCT-1999; 990S-0159431;
PR 14-OCT-1999; 990S-0159637;
PR 18-OCT-1999; 990S-0159638;
PR 21-OCT-1999; 990S-0160741;
PR 21-OCT-1999; 990S-0160747;
PR 21-OCT-1999; 990S-0160768;
PR 21-OCT-1999; 990S-0160770;
PR 21-OCT-1999; 990S-0160814;
PR 21-OCT-1999; 990S-0160815;
PR 22-OCT-1999; 990S-0160980;
PR 22-OCT-1999; 990S-0160981;
PR 25-OCT-1999; 990S-0161404;
PR 25-OCT-1999; 990S-0161405;
PR 25-OCT-1999; 990S-0161406;
PR 26-OCT-1999; 990S-0161456;
PR 26-OCT-1999; 990S-0161460;
PR 26-OCT-1999; 990S-0161461;
PR 28-OCT-1999; 990S-0161920;
PR 28-OCT-1999; 990S-0161922;
PR 28-OCT-1999; 990S-0161943;
PR 29-OCT-1999; 990S-0162142;
PR 29-OCT-1999; 990S-0162142;

Query Match 92.9%; Score 13; Exp 21; Length 1105;
Best local similarity 100.0%; Pred. No. 4; Gap 02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 characterate 14
DB 24 characterate 242

RESULT 7
ID C76482/c
XX C76482 standard: cDNA; 1864 BP.
XX C76482.
XX C76482.
DT 08-FEB-2001 (first entry)

```

```

XX Human ORFX ORF2037 polynucleotide sequence SEQ ID NO:4074.
DE Human, open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW valvular; antipsoriatic; antiparkinsonian; neurotrophic; neurotrophic;
KW anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antineumatic; antithyroid;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester; storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombolysis; contraceptive; ss.
XX Homo sapiens.
OS Homo sapiens.
XX W0200059473-A2.
XX 05 OCT 2000.
XX 31 MAR 2000; 2000W: 9908521.
XX 31 MAR-1999; 990S-0127607.
XX 02-APR-1999; 990S-0127636.
XX 05-APR-1999; 990S-0127728.
XX 30-MAR-2000; 2000MS-0540763.
XX (CURA-) CURAGEN CORP.
XX Shimkets RA, Leuch M;
XX WPI: 2000 65242/57.
XX P-PSDB; 842273.
XX Nucleotide acids and peptides derived from open reading frame X.
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease.
XX Claim 5; Page 3262 3263; 557pp; English.
XX C74446 to C77636 encode the proteins given in 840247 to 843347, which
XX represent the human ORFX open reading frames 1 to 461. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; valvular;
XX antipsoriatic; antiparkinsonian; neurotrophic; neuroprotective; osteopathic;
XX anticonvulsant; antiarthritis; immunosuppressant; immunostimulant;
XX cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
XX dermatological; immunosuppressive; antiinflammatory; antibacterial;
XX antiviral; antifungal; antineumatic; antithyroid; and antianemic. The
XX sequences can be used for determining the presence of or predisposition
XX to, or preventing or treating pathological conditions associated with an
XX ORFX associated disorder, the nucleic acids can be used to express ORFX
XX proteins in gene therapy vectors. The proteins and nucleic acids may be
XX used to treat cancers, proliferative disorders, neurodegenerative
XX disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
XX diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
XX storage, systemic lupus erythematosus, severe combined immunodeficiency
XX (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
XX disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
XX cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
XX enhance coagulation, to inhibit thrombolysis, and as a contraceptive
XX
SO Sequence 1864 BP; 511 A; 444 C; 451 G; 466 T; 2 Other;

```

Query Match 92.9%; Score 13; Exp 21; Length 1864;
 Best local similarity 100.0%; Pred. No. 4; Gap 02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 characterate 14


```

PN W0200046364-A1.
XX
PD 10-AUG-2000.
XX
PF 03-FEB-2000; 2000W0 US02867.
XX
PR 05-FEB-1999; 99US-0118880.
XX
PA (GABR ) GABR RES PABD0.
XX
PI Daniel DL, Prasadgith VM, Emiliyev T, Douglaz JT.
XX
DE WPI: 2000-506070/45.
XX
PT Propagation of adenovirus in a host cell, comprising expressing an
PT artificial recombinant receptor on the surface of a host cell and
PT infecting the host with the adenovirus, useful for cell specific gene
PT therapy -
XX
PS Example 8; Page 21; 54pp; English.
XX
CC The present invention relates to a method for propagating adenoviruses
CC in host cells. The method comprises expressing an artificial recombinant
CC receptor on the surface of a host cell and infecting the host with the
CC adenovirus. The virus binding domain of the novel receptor consists of
CC adenovirus serotype 5 (Ad5) fiber protein with a C-terminal His tag
CC (fiber-6His). The present sequence is a coding sequence for fiber-6His
CC peptide. The present sequence was used in the production of vectors
CC encoding the novel receptor of the present invention. The novel vectors
CC are useful for cell specific gene therapy. See B15187 for details of the
CC poly-His tag used to construct the fiber-6His peptide.
SQ Sequence 81 BP; 16 A; 12 C; 42 G; 21 T; 0 other;

Query Match: 90.6%; Score 12.4; BP 21; Length 91;
Best Local Similarity 92.9%; Pred. No. 9.7e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Q7 1 tetraacetate 14
II 78 TTAGCCCTCCGCAIC 65

RESULT 15
C51994/
II C51994 standard; cDNA; 84 BP.
XX
A5 C51994:
XX
PI 06-0CT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID No: 36069.
XX
XX
FW human; 5' EST; expressed sequence tag, secreted protein, cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
CS Homo sapiens.
XX
IN EP1034401-A2.
XX
PC 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST ) GENSEI.
XX
PI Thomas Milne Edwards J, Duclert A, Giordano J;
XX
DE WPI: 2000-500381/45.
XX

```

```

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic RNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 36069; 71pp - CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic RNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 84 BP; 25 A; 18 C; 46 G; 2 T; 3 other;

Query Match: 98.6%; Score 12.4; BP 21; Length 84;
Best Local Similarity 92.9%; Pred. No. 9.7e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Q7 1 tetraacetate 14
II 44 TCTGCCCTCCGCAIC 31
DB

```

Search completed: May 1, 2001, 04:00:19
Job time: 7124 sec



Copyright (c) 1993 - 2000 Computer Ltd.

GenScore version 4.5

OR nucleic nucleic search, using SW model

Run on: May 1, 2001, 09:03:50, Search time: 2512.88 Seconds
(without alignments)
82.165 Million cell updates/sec

Index: us_09_41_829a_4_copy_1_14
Database: 14
1 request/cycle 14

Scoring table: IDENTITY_NUC
Gapop: 10,0 / Gapext: 1,0

Searched: 1204235 seqs, 74792952 residues

Total number of hits satisfying chosen parameters: 2566470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match 0%

Listing first 45 summaries

Database: GenBank: *

| | | |
|-----|----------|---|
| 1: | ab_ba1: | * |
| 2: | ab_ba2: | * |
| 3: | ab_ba3: | * |
| 4: | ab_ba4: | * |
| 5: | ab_ba5: | * |
| 6: | ab_ba6: | * |
| 7: | ab_ba7: | * |
| 8: | ab_ba8: | * |
| 9: | ab_ba9: | * |
| 10: | ab_ba10: | * |
| 11: | ab_ba11: | * |
| 12: | ab_ba12: | * |
| 13: | ab_ba13: | * |
| 14: | ab_ba14: | * |
| 15: | ab_ba15: | * |
| 16: | ab_ba16: | * |
| 17: | ab_ba17: | * |
| 18: | ab_ba18: | * |
| 19: | ab_ba19: | * |
| 20: | ab_ba20: | * |
| 21: | ab_ba21: | * |
| 22: | ab_ba22: | * |
| 23: | ab_ba23: | * |
| 24: | ab_ba24: | * |
| 25: | ab_ba25: | * |
| 26: | ab_ba26: | * |
| 27: | ab_ba27: | * |
| 28: | ab_ba28: | * |
| 29: | ab_ba29: | * |
| 30: | ab_ba30: | * |
| 31: | ab_ba31: | * |
| 32: | ab_ba32: | * |
| 33: | ab_ba33: | * |
| 34: | ab_ba34: | * |
| 35: | ab_ba35: | * |
| 36: | ab_ba36: | * |
| 37: | ab_ba37: | * |
| 38: | ab_ba38: | * |
| 39: | ab_ba39: | * |
| 40: | ab_ba40: | * |
| 41: | ab_ba41: | * |
| 42: | ab_ba42: | * |
| 43: | ab_ba43: | * |

44: em_ov: *

45: em_par: *

46: em_ph: *

47: em_pl: *

48: em_ro: *

49: em_ss: *

50: em_sy: *

51: em_un: *

52: em_v1: *

53: ab_ss1: *

54: ab_ss2: *

55: ab_ss3: *

56: ab_sy: *

57: ab_un: *

58: ab_v11: *

59: ab_v12: *

60: ab_bt1: *

61: ab_bt2: *

62: ab_bt3: *

63: ab_bt4: *

64: ab_bt5: *

65: ab_bt6: *

66: ab_bt7: *

67: ab_bt8: *

68: ab_bt9: *

69: ab_bt10: *

70: ab_bt11: *

71: ab_bt12: *

72: ab_bt13: *

73: ab_bt14: *

74: ab_bt15: *

75: ab_bt16: *

76: ab_bt17: *

77: ab_bt18: *

78: ab_bt19: *

79: ab_bt20: *

80: ab_bt21: *

81: ab_bt22: *

82: ab_bt23: *

83: ab_bt24: *

84: ab_bt25: *

85: ab_pr1: *

86: ab_pr2: *

87: ab_pr3: *

88: ab_pr4: *

89: ab_pr5: *

90: ab_pr6: *

91: ab_pr7: *

92: ab_pr8: *

93: ab_pr9: *

94: ab_pr10: *

95: ab_pr11: *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 14 | 100.0 | 224 | 13 | AY017079 oryza sat |
| 2 | 14 | 100.0 | 708 | 72 | AC049025 Glardia i |
| 3 | 14 | 100.0 | 795 | 74 | AC066869 Glardia i |
| 4 | 14 | 100.0 | 940 | 72 | AC049024 Glardia i |
| 5 | 14 | 100.0 | 1002 | 9 | AR042535 Sequence |
| 6 | 14 | 100.0 | 1002 | 92 | HSA223040 Homo Sapi |
| 7 | 14 | 100.0 | 2649 | 92 | HSA275977 Homo Sapi |
| 8 | 14 | 100.0 | 5081 | 8 | AB012720 oncorhine |
| 9 | 14 | 100.0 | 3245 | 91 | H08223094 Homo Sapi |
| 10 | 14 | 100.0 | 9859 | 94 | M08F082 |
| 11 | 14 | 100.0 | 45565 | 8 | AF151319 Mus musculu |
| | | | | | AL021541 Fugu rubr |

```

12 14 100.0 46738 91 HS594110
13 14 100.0 57198 66 AC020115
14 14 100.0 62748 4 AC005417
15 14 100.0 63288 72 AC011762
16 14 100.0 68338 68 AC017440
17 14 100.0 83218 90 AL559880
18 14 100.0 89479 61 AC010444
19 14 100.0 103547 61 AC010009
20 14 100.0 110184 87 AC020906
21 14 100.0 111372 79 AL53145
22 14 100.0 112874 78 AF277415
23 14 100.0 112874 78 AF277415
24 14 100.0 120507 89 AL136358
25 14 100.0 133080 61 AC009523
26 14 100.0 133090 65 AC017815
27 14 100.0 135276 75 AC073603
28 14 100.0 148184 65 AC017777
29 14 100.0 150069 72 AC041042
30 14 100.0 151507 76 AC078802
31 14 100.0 157757 80 AL455344
32 14 100.0 161933 59 AC024704
33 14 100.0 164918 57 AC021599
34 14 100.0 164918 57 AC021599
35 14 100.0 168312 60 AC007484
36 14 100.0 171472 60 AC007400
37 14 100.0 175793 60 AC007401
38 14 100.0 176775 65 AC016978
39 14 100.0 179323 62 AC011852
40 14 100.0 179937 87 AC016683
41 14 100.0 180521 73 AC057817
42 14 100.0 184092 65 AC018715
43 14 100.0 191161 64 AC016043
44 14 100.0 198316 66 AC019157
45 14 100.0 224744 57 AC022165
14 100.0 265512 80 AL455388

ALIGNMENTS

RESULT 1
LOCUS AY017679 224 bp DNA PLN 07-FEB-2001
DEFINITION Oryza sativa microsatellite M809094 containing (AC)X12, closest to marker P2443, genomic sequence.
ACCESSION AY017679
VERSION AY017679.1 GI:12700721
KEYWORDS Oryza sativa.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryote; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzoae; Oryza.
REFERENCE 1 (bases 1 to 224)
AUTHORS Tao, N., Barbazuk, W.R., Liu, J., Wu, K., and Barry, G.F.
TITLE Simple sequence repeats from Monsanto rice genomic sequences
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 224)
AUTHORS Tao, N., Barbazuk, W.R., Liu, J., Wu, K., and Barry, G.F.
TITLE Direct Submission
JOURNAL Submitted (10-JAN-2001) Genomics, Monsanto, 800 North Lindbergh Blvd., Creve Coeur, MO 63167, USA
COMMENT Derived from rice genomic sequences generated from the Monsanto Rice Genome Sequencing project. Please see http://www.rice-research.org for more information. The sequence data were produced primarily in the laboratories of Dr. Leroy Hood at the University of Washington in Seattle.
FEATURES
source
location/Qualifiers
1..224
organism="Oryza sativa"
/db_xref="X000453"
repeat_region 1..224
note="microsatellite M809094"
/rpt_type=tandem

```

```

BASE COUNT 53 a 80 c 52 q 39 t
ORIGIN
Query Match 100.0% Score 14; DB 13; Length 224;
Best Local Similarity 100.0% Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tctgctccgcatc 14
|||||
Db 78 tctgctccgcatc 91

RESULT 2
AC049025/c
LOCUS AC049025 708 bp DNA HTG 14-APR-2000
DEFINITION Giardia intestinalis clone KJ1363 strain WB-C6, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC049025
VERSION AC049025.1 GI:7554165
KEYWORDS HTG; PHASE2.
SOURCE Giardia intestinalis.
ORGANISM Giardia intestinalis
Eukaryote; Diplomonadida; Hexamitidae; Giardia.
REFERENCE 1 (bases 1 to 708)
AUTHORS Morrison, H.G., McArthur, A.G., Nixon, J., Eskin, N.Q., Kim, D., Crocker, M.C., Hinkle, G., Holder, M.E. and Sogin, M.L.
TITLE Giardia: a model for ancient eukaryotic genome analysis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 708)
AUTHORS Kim, D., Morrison, H.G., McArthur, A.G., Nixon, J., Eskin, N.Q., Crocker, M.C., Hinkle, G., Holder, M.E. and Sogin, M.L.
TITLE Direct Submission
JOURNAL Submitted (14-APR-2000) Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA 02543-1015, USA
COMMENT * NOTE: This record contains 1 individual
* sequencing reads that have not been assembled into
* contigs. Rups of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved. 708: contig of 708 bp in length.
FEATURES
source
location/Qualifiers
1..708
organism="Giardia intestinalis"
/strain="WB-C6"
/db_xref="taxon:5741"
/clone="KJ1363"
BASE COUNT 217 a 149 c 159 q 187 t 7 others
ORIGIN
Query Match 100.0% Score 14; DB 72; Length 708;
Best Local Similarity 100.0% Pred. No. 1.2e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tctgctccgcatc 14
|||||
Db 580 tctgctccgcatc 567

RESULT 3
LOCUS AC06669 795 bp DNA HTG 25-APR-2000
DEFINITION Giardia intestinalis clone N10870 strain WB-C6, LOW-PASS SEQUENCE SAMPLING.

```


ACCESSION A04624
 VERSION A04624.1 GI:2645609
 KEYWORDS HIGS PHASED
 SOURCE Giardia intestinalis
 ORGANISM Giardia intestinalis
 Eukaryote; Diplomonadida; Hexamitidae; Giardinae; Giardia
 1 (bases 1 to 795)
 AUTHOR Crickor, M.C., Hunkle, G., Holder, M.E., and Sodini, M.L.
 TITLE Giardia: a model for ancient eukaryotic genome analysis
 JOURNAL unpublished
 REFERENCE 2 (bases 1 to 795)
 AUTHOR Nixon, J., Morrison, J.G., McArthur, A.G., Park, N.G., Kim, B., Crickor, M.C., Hunkle, G., Holder, M.E., and Sodini, M.L.
 TITLE Direct Submission
 JOURNAL Molecular Biology and Evolution, Marine Biotechnological Laboratory, 7 MBL Street, Woods Hole, MA 02543-1015, USA
 COMMENT
 * NOTE: This record contains 1 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low pass sequence sampling is useful for
 * identifying clones that may be gene rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.
 * 1 795: contig of 795 bp in length.
 * location/Qualifiers
 * 1..795
 * Zorganism "Giardia intestinalis"
 * ZSTRID "WR-06"
 * ZR_Strain "K10878"
 * ZLab "MBL"

Query Match 100.0% Score 14; DB 14; Length 795;
 Best Local Similarity 100.0%; Pred. No. 1,10003;
 Matches 14; Conservative 0; Mismatches 0; Gaps 0;
 QY 1 tetanocysteate 14
 DB 106 TCGCTCCGAC 93

FEATURES
 source
 location/Qualifiers
 1..795
 Zorganism "Giardia intestinalis"
 ZSTRID "WR-06"
 ZR_Strain "K10878"
 ZLab "MBL"

BASE COUNT 227 a 168 c 181 g 228 t 1 others
 ORIGIN 1..795: contig of 795 bp in length.
 1..795
 Zorganism "Giardia intestinalis"
 ZSTRID "WR-06"
 ZR_Strain "K10878"
 ZLab "MBL"

Query Match 100.0% Score 14; DB 14; Length 795;
 Best Local Similarity 100.0%; Pred. No. 1,10003;
 Matches 14; Conservative 0; Mismatches 0; Gaps 0;
 QY 1 tetanocysteate 14
 DB 106 TCGCTCCGAC 93

RESULT 4
 LOCUS A04624 940 bp DNA 14-APR-2000
 DEFINITION Giardia intestinalis clone K10878 strain WR-06, LOW PASS SEQUENCE
 AUTHORS Crickor, M.C., Hunkle, G., Holder, M.E., and Sodini, M.L.
 TITLE Giardia: a model for ancient eukaryotic genome analysis
 JOURNAL unpublished
 REFERENCE 2 (bases 1 to 950)
 AUTHOR Kim, J., Morrison, J.G., McArthur, A.G., Park, N.G., Kim, B., Crickor, M.C., Hunkle, G., Holder, M.E., and Sodini, M.L.
 TITLE Direct Submission
 JOURNAL Molecular Biology and Evolution, Marine Biotechnological Laboratory, 7 MBL Street, Woods Hole, MA 02543-1015, USA
 COMMENT
 * NOTE: This record contains 1 individual

* sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low pass sequence sampling is useful for
 * identifying clones that may be gene rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.
 * 1 940: contig of 940 bp in length.
 * location/Qualifiers
 * 1..940
 * Zorganism "Giardia intestinalis"
 * ZSTRID "WR-06"
 * ZR_Strain "K10878"
 * ZLab "MBL"

BASE COUNT 287 a 190 c 201 g 252 t
 ORIGIN 1..940: contig of 940 bp in length.
 1..940
 Zorganism "Giardia intestinalis"
 ZSTRID "WR-06"
 ZR_Strain "K10878"
 ZLab "MBL"

Query Match 100.0% Score 14; DB 14; Length 940;
 Best Local Similarity 100.0%; Pred. No. 1,10003;
 Matches 14; Conservative 0; Mismatches 0; Gaps 0;
 QY 1 tetanocysteate 14
 DB 578 TCGCTCCGAC 965

RESULT 5
 LOCUS A042535 1002 bp DNA 29-SEP-1998
 DEFINITION Sequence 4 from patient US 5811519.
 AUTHORS Lee, B., Lucas, S., de Smidt, C., Godelaine, D., and Boon, E.
 TITLE LL-1 tumor specific genes
 JOURNAL Patent: US 5811519-A 22 SEP-1998
 FEATURES
 location/Qualifiers
 1..1002
 source
 BASE COUNT 179 a 302 c 336 g 145 t
 ORIGIN 1..1002: contig of 1002 bp in length.
 1..1002
 Zorganism "unknown"
 ZSTRID "WR-06"
 ZR_Strain "K10878"
 ZLab "MBL"

Query Match 100.0% Score 14; DB 14; Length 1002;
 Best Local Similarity 100.0%; Pred. No. 1,10003;
 Matches 14; Conservative 0; Mismatches 0; Gaps 0;
 QY 1 tetanocysteate 14
 DB 1 TCGCTCCGAC 14

RESULT 6
 LOCUS HSA223040 1002 bp mRNA 24 JUN 1998
 DEFINITION Homo sapiens mRNA for LAGE-1b protein.
 AUTHORS Hunkle, M.C., Hunkle, G., Holder, M.E., and Sodini, M.L.
 TITLE Giardia: a model for ancient eukaryotic genome analysis
 JOURNAL unpublished
 REFERENCE 2 (bases 1 to 950)
 AUTHOR Kim, J., Morrison, J.G., McArthur, A.G., Park, N.G., Kim, B., Crickor, M.C., Hunkle, G., Holder, M.E., and Sodini, M.L.
 TITLE Direct Submission
 JOURNAL Molecular Biology and Evolution, Marine Biotechnological Laboratory, 7 MBL Street, Woods Hole, MA 02543-1015, USA
 COMMENT
 * NOTE: This record contains 1 individual

For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.trinity-ed.com/sequence/>) or send email to bdp@trinity.berkeley.edu.

PI library location: 26-11.

FEATURES

source
1..62748
location/qualifiers
1..62748
/organism="Drosophila melanogaster"
/strain="2; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2L"
/map="29F1-29F3"
/clone="115 "PI library, partial SAGE in pMS582(ect14AD10)"
/feature="F1:582111 (5566)"
/note "This clone is a partial bridge extending from bp 1 minimally overlapping its distal neighbor pS00901 (5255) to PI end at bp 62,749."

BASE COUNT 19254 a 12984 c 12547 g 17963 t

ORIGIN

Query Match 100.0%; Score 14; DB 4; Length 62748;
Best Local Similarity 100.0%; Prod. No. 5,46,02;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tetanactetate 14
IIIIIIIIIIII
DB 56376 TCTGCTTCGATC 56389
AC011762 63288 bp DNA HTG 31-JAN-2000
Drosophila melanogaster chromosome 3 clone WAC84901 (D1151)
Recl-98 49.8.1 map 100E-100F strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***, 65 unordered pieces.

ACCESSION

AC011762.2 GI:6838818

VERSION

HTG: HUGS-PHASE1

KEYWORDS

fruit fly

SOURCE

Drosophila melanogaster

ORGANISM

Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 63288)
Celisner, S. E., Akayama, A., Aranda, T. L., Baxter, E., Blacraj, R. G., Rutenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L., Boyle, C. M., Farfan, D. E., Gallo, R., George, R. A., Harris, N. L., Hinko, A., Hopkins, P. A., Johnston, K. A., Hummasti, S. P., Karia, K., Kearney, L., Lee, B., Lewis, S., Li, P., Lind, B., Moshure, J. A., Roshini, M., Nixon, K., Facib, J. M., Park, S., Pfeiffer, B., Richards, S., Sothi, H., Swirskas, P., Wan, P. H., Webster, D., Woolley, P., Yang, S., Yee, M., Yu, C. and Rubin, G. M.

REFERENCE

Unpublished

2 (bases 1 to 63288)

Celisner, S. E., Akayama, A., Aranda, T. L., Baxter, E., Blacraj, R., Rutenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L., Boyle, C. M., Farfan, D. E., Gallo, R., George, R. A., Harris, N. L., Hopkins, P. A., Johnston, K. A., Hummasti, S. P., Karia, K., Kearney, L., Kim, E., Lee, B., Lewis, S., Li, P., Lomolan, M., Mada, P., Moshure, J. A., Roshini, M., Nixon, K., Facib, J. M., Park, S., Pfeiffer, B., Roshini, M., Swirskas, P., Wan, P. H., Webster, D., Woolley, P., Yang, S., Yee, M., Yu, C. and Rubin, G. M.

TITLE

Submitted

JOURNAL

Submitted (14-sep-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA

COMMENT

on Jan 31, 2000 this sequence version replaced 41694705.
For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.trinity-ed.com/sequence/>) or send email to bdp@trinity.berkeley.edu. All contigs in this submission were

the following cutoffs: length >= 200 bases.
* NOTE: This is a 'working draft' sequence. It currently consists of 65 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1
688: contig of 608 bp in length
689: gap of unknown length
1481: contig of 793 bp in length
1561: gap of unknown length
2344: contig of 703 bp in length
2345: gap of unknown length
2943: contig of 599 bp in length
3023: gap of unknown length
3511: contig of 487 bp in length
3590: gap of unknown length
4001: contig of 411 bp in length
4081: gap of unknown length
4539: contig of 449 bp in length
4610: gap of unknown length
5268: contig of 658 bp in length
5348: gap of unknown length
5689: contig of 341 bp in length
5759: gap of unknown length
6775: contig of 1006 bp in length
6855: gap of unknown length
7267: contig of 412 bp in length
7347: gap of unknown length
8033: contig of 586 bp in length
8113: gap of unknown length
9116: contig of 1003 bp in length
9196: gap of unknown length
9942: contig of 746 bp in length
10023: gap of unknown length
10608: contig of 586 bp in length
10688: gap of unknown length
11131: contig of 443 bp in length
12211: gap of unknown length
11790: contig of 579 bp in length
11870: gap of unknown length
12565: contig of 695 bp in length
12645: gap of unknown length
13414: contig of 764 bp in length
13494: gap of unknown length
14368: contig of 874 bp in length
14418: gap of unknown length
15088: contig of 640 bp in length
15168: gap of unknown length
16251: contig of 1083 bp in length
16331: gap of unknown length
17722: contig of 1441 bp in length
17852: gap of unknown length
18444: contig of 592 bp in length
18523: gap of unknown length
19607: contig of 1083 bp in length
19687: gap of unknown length
20313: contig of 626 bp in length
20393: gap of unknown length
21212: contig of 819 bp in length
21292: gap of unknown length
22506: contig of 1214 bp in length
22596: gap of unknown length
23231: contig of 635 bp in length
23301: gap of unknown length
24809: contig of 1398 bp in length
24889: gap of unknown length
25632: contig of 743 bp in length
25712: gap of unknown length
27286: contig of 1674 bp in length
27366: gap of unknown length


```

2767 28519: contig of 1154 bp in length
2852 28599: gap of unknown length
2860 30457: contig of 1858 bp in length
3046 30547: gap of unknown length
3058 42147: contig of 1600 bp in length
4215 42217: gap of unknown length
4220 44026: contig of 1809 bp in length
4403 44106: gap of unknown length
4410 46269: contig of 2163 bp in length
4627 46349: gap of unknown length
4635 48117: contig of 2368 bp in length
4812 48797: gap of unknown length
4879 48957: contig of 1699 bp in length
4896 49576: gap of unknown length
4958 49774: contig of 2197 bp in length
4979 49987: gap of unknown length
4999 44067: contig of 1144 bp in length
4407 44578: gap of unknown length
4458 47758: contig of 3691 bp in length
4776 47848: gap of unknown length
4785 48751: contig of 913 bp in length
4876 48831: gap of unknown length
4883 49498: contig of 667 bp in length
4949 49578: gap of unknown length
4957 50322: contig of 744 bp in length
5033 50402: gap of unknown length
5040 50955: contig of 553 bp in length
5096 51035: gap of unknown length
5103 51540: contig of 495 bp in length
5154 51610: gap of unknown length
5161 52241: contig of 631 bp in length
5222 52321: gap of unknown length
5232 52864: contig of 543 bp in length
5286 52944: gap of unknown length
5294 53398: contig of 454 bp in length
5339 53478: gap of unknown length
5347 54255: contig of 777 bp in length
5426 54335: gap of unknown length
5433 54890: contig of 555 bp in length
5489 54970: gap of unknown length
5497 55570: contig of 600 bp in length
5557 55650: gap of unknown length
5565 56217: contig of 567 bp in length
5621 56297: gap of unknown length
5629 56880: contig of 583 bp in length
5688 56960: gap of unknown length
5696 57540: contig of 576 bp in length
5754 57616: gap of unknown length
5761 58114: contig of 498 bp in length
5811 58194: gap of unknown length
5819 58760: contig of 566 bp in length
5876 58840: gap of unknown length
5884 59129: contig of 289 bp in length
5912 59209: gap of unknown length
5920 59792: contig of 583 bp in length
5979 59872: gap of unknown length
5987 60440: contig of 568 bp in length
6044 60520: gap of unknown length
6052 61168: contig of 648 bp in length
6116 61248: gap of unknown length
6124 61795: contig of 547 bp in length
6179 61875: gap of unknown length
6187 62431: contig of 556 bp in length
6243 62511: gap of unknown length
6251 63288: contig of 777 bp in length.

```

FEATURES
source

```

/contig/qualifiers
1..36298
/contig "Brosophila melanogaster"
/strain "9: cn bw sp"
/feature "1..1656..223"
/contig "9"
/contig "100E:100F"
/clone "BACR49B01 (D1151) RPEC1 98 49.B.1"

```

```

/clone_lib "RPEC1 98 (Roswell Park Cancer Institute)
Brosophila melanogaster BAC library, partial Ecoli 10.
pBACE3.6"
BASE COUNT 16122 a 14174 c 12761 g 16106 t 5124 others
ORIGIN
Query Match 100.0%; Score 14; LR 62; length 63288;
Best Local Similarity 100.0%; Pred. No. 5,4002;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
cy 1 tetra nucleotide 14
lib 18787 TCTGCTCTCCGATC 18774
Search completed: May 1, 2001, 04:52:12
Job time: 11665 sec

```


Copyright (c) 1994 - 2000 CompuLink Ltd.

CM nucleotide nucleotide search using sw model

Run on: May 1, 2001, 04:55:25 : Search time 156.99 seconds
(without alignments)
16,189 Million cell updates/sec

File: US-09-341-829A_4_copy_1_14
Perfect score: 14
Sequence: 14

Scoring table: IDENTITY_80%
Gap: 10, 0, GapExt: 1, 0

Searched: 602621 seqs, 8740144 residues

Total number of hits satisfying chosen parameters: 605242

Minimum DB seq length: 6
Maximum DB seq length: 2000000

Post processing: Minimum Match 0%
Maximum Match 100%
Using first 4% summaries

Database: ISSUED_PATENT_S_NA:*

| Seq | Score | Match | Length | ID | Description |
|-----|-------|-------|--------|----|--------------------|
| 1 | 14.0 | 100.0 | 1002 | 1 | US-08-791-495-4 |
| 2 | 12.4 | 88.6 | 20 | 2 | US-08-457-2746-15 |
| 3 | 12.4 | 88.6 | 24 | 3 | US-09-011-669-2 |
| 4 | 12.4 | 88.6 | 51 | 5 | US-08-085-0474-5 |
| 5 | 12.4 | 88.6 | 51 | 5 | US-08-085-0474-14 |
| 6 | 12.4 | 88.6 | 71 | 2 | US-08-894-578-219 |
| 7 | 12.4 | 88.6 | 550 | 1 | US-08-248-016-11 |
| 8 | 12.4 | 88.6 | 550 | 1 | US-08-451-601-11 |
| 9 | 12.4 | 88.6 | 550 | 5 | US-08-06761-11 |
| 10 | 12.4 | 88.6 | 519 | 2 | US-08-611-757-48 |
| 11 | 12.4 | 88.6 | 519 | 5 | US-08-0595-0598-48 |
| 12 | 12.4 | 88.6 | 768 | 3 | US-08-946-026-42 |
| 13 | 12.4 | 88.6 | 858 | 4 | US-08-475-416A-14 |
| 14 | 12.4 | 88.6 | 1371 | 2 | US-08-919-741-1 |
| 15 | 12.4 | 88.6 | 1371 | 2 | US-08-919-741-7 |
| 16 | 12.4 | 88.6 | 1371 | 2 | US-08-795-395-1 |
| 17 | 12.4 | 88.6 | 1479 | 2 | US-08-199-485-1 |
| 18 | 12.4 | 88.6 | 1569 | 2 | US-08-368-260-1 |
| 19 | 12.4 | 88.6 | 1569 | 2 | US-08-559-221-1 |
| 20 | 12.4 | 88.6 | 1669 | 3 | US-08-860-519-11 |
| 21 | 12.4 | 88.6 | 1685 | 6 | US-09-034-1 |
| 22 | 12.4 | 88.6 | 1784 | 4 | US-09-226-741-2 |
| 23 | 12.4 | 88.6 | 1872 | 1 | US-08-428-925-1 |
| 24 | 12.4 | 88.6 | 1872 | 1 | US-08-435-444-4 |
| 25 | 12.4 | 88.6 | 1872 | 1 | US-08-435-446-4 |
| 26 | 12.4 | 88.6 | 1872 | 1 | US-08-428-927-1 |
| 27 | 12.4 | 88.6 | 1872 | 1 | US-08-428-298-1 |

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

US-08-791-495-4
Sequence 4, Application US/08791495
Patent No. 5811519
GENERAL INFORMATION:
APPLICANT: Leth, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: Le Smet, Charles
ATTORNEY: J. B. Le Smet, Advocate
AGENCY: Bioscience Resource Project
TITLE: INVENTION FOR A NEW SUPERGENES
NUMBER OF SEQUENCES: 14
ADDRESS: Well, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08791495
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
REFERENCE/BOOK NUMBER: 104617005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-2441
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1002 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MODE: TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CUS
LOCATION: 65..697
US-08-791-495-4

ALIGNMENTS

RESULT 1
US-08-791-495-4
Sequence 4, Application US/08791495
Patent No. 5811519
GENERAL INFORMATION:
APPLICANT: Leth, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: Le Smet, Charles
ATTORNEY: J. B. Le Smet, Advocate
AGENCY: Bioscience Resource Project
TITLE: INVENTION FOR A NEW SUPERGENES
NUMBER OF SEQUENCES: 14
ADDRESS: Well, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08791495
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
REFERENCE/BOOK NUMBER: 104617005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-2441
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1002 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MODE: TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CUS
LOCATION: 65..697
US-08-791-495-4

Query Match 100.0%; Score 14; DB 1; Length 1002;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tetagctctccatc 14
 DB 1 TCGTCGTCGTCGTC 14

RESULT 2
 US-08-457-273B-15/c
 : Sequence 15, Application US/09045777ap
 : Patent No. 5849995
 : GENERAL INFORMATION:
 : APPLICANT: Hayden, Michael
 : APPLICANT: Lin, Biagand
 : APPLICANT: Nasir, Jamal
 : TITLE OF INVENTION: Mouse Model for Huntington's Disease and
 : TITLE OF INVENTION: Related DNA Sequences
 : NUMBER OF SEQUENCES: 42
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Virginia Bennett
 : STREET: PO Box 37428
 : CITY: Raleigh
 : STATE: No. 5849995th Carolina
 : COUNTRY: US
 : ZIP: 27627
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: 08/09045777ap
 : FILING DATE:
 : CLASSIFICATION: 800
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Bennett, Virginia C.
 : REGISTRATION NUMBER: 37,092
 : REFERENCE/DOCKET NUMBER: 3477-85A
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 919-854-1400
 : TELEFAX: 919-854-1401
 : INFORMATION FOR SEQ ID NO: 15:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 20 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: cDNA
 : US-08-457-273B-15

Query Match 88.6%; Score 12.4; DB 2; Length 20;
 Best Local Similarity 92.9%; Pred. No. 2.8e+02;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tetagctctccatc 14
 DB 15 TCGTCGTCGTCGTC 2

RESULT 5
 US-09-011-600-2/c
 : Sequence 2, Application US/09011600
 : Patent No. 6184954
 : GENERAL INFORMATION:
 : APPLICANT: BAR-SHAVIT, RAHIEL
 : TITLE OF INVENTION: A METHOD AND KIT FOR EVALUATING THE
 : TITLE OF INVENTION: METASTATIC TENDENCY OF TUMORS
 : NUMBER OF SEQUENCES: 5

: CORRESPONDENCE ADDRESS:
 : ADDRESSEE: OFFICE & REPAIRING, PLC
 : STREET: P.O. BOX 19928
 : CITY: ALEXANDRIA
 : STATE: VA
 : COUNTRY: USA
 : ZIP: 22320
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09011600
 : FILING DATE: 10-MAR-1998
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: BERRIDGE, WILLIAM P
 : REGISTRATION NUMBER: 30,024
 : REFERENCE/DOCKET NUMBER: JAO 40800
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (703)-836-6400
 : TELEFAX: (703)-836-2787
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 24 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: DNA (genomic)
 : US-09-011-600-2

Query Match 88.6%; Score 12.4; DB 4; Length 24;
 Best Local Similarity 92.9%; Pred. No. 2.8e+02;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tetagctctccatc 14
 DB 18 TCGTCGTCGTCGTC 5

RESULT 4
 PCT-US95-03742-5
 : Sequence 5, Application EC/TUS9503742
 : GENERAL INFORMATION:
 : APPLICANT: The IAR Research Foundation
 : TITLE OF INVENTION: Ligands Added to Adenovirus Fiber
 : NUMBER OF SEQUENCES: 18
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Patricia L. Pabst
 : STREET: 2800 One Atlantic Center
 : STREET: 1201 West Peachtree Street
 : CITY: Atlanta
 : STATE: Georgia
 : COUNTRY: USA
 : ZIP: 30309-1450
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: PCT/US95/03742
 : FILING DATE:
 : CLASSIFICATION:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Pabst, Patricia L.
 : REGISTRATION NUMBER: 31,284
 : REFERENCE/DOCKET NUMBER: 1G1101
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (404) 873-8794
 : TELEFAX: (404) 873-8795

```

1  INFORMATION FOR SEQ ID NO: 5:
2  SEQUENCE CHARACTERISTICS:
3  LENGTH: 51 base pairs
4  TYPE: nucleic acid
5  STRANDEDNESS: single
6  topology: linear
7  MOLECULE TYPE: DNA
8  HYDROPHOBIC: NO
9  ANTI-SENSE: NO
10 FEATURE:
11 NAME/KEY: misc feature
12 LOCATION: 1..51
13 OTHER INFORMATION: cloning the linker peptide of Sequence ID No. 1."
14 US95/03742.5

Query Match: 88.4%; Score 12.4; DB %: Length 51;
Best Local Similarity: 92.9%; Prod. No. 2.8e+02;
Matches: 14; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;

QY 1 tetanotoxin 14
DB 14 TCAGGCTGGGATC 26

RESULT 5
US-09-341-829a-4_14/0
US-09-341-829a-4_14/0
Patent No. 5998142
GENERAL INFORMATION:
APPLICANT: GOLD et al.
TITLE OF INVENTION: SYSTEMATIC EVOLUTION OF LIBRARIES
BY EXPONENTIAL ENRICHMENT:
TITLE OF INVENTION: CHEM-SELEX
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratsch, L.L.C.
STREET: 8400 E. Front Street, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/341,829a-4_14/0
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 5998142
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,141
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/546,428
FILING DATE: 11-JUNE-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/117,991
FILING DATE: 9-SEPTEMBER-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/124,945
FILING DATE: 17-SEPTEMBER-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/199,597
FILING DATE: 22-FEBRUARY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/244,997
FILING DATE: 28-APRIL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/309,245
FILING DATE: 20-SEPTEMBER-1994
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 44,215
REFERENCE: 5998142
TELEPHONE: (303) 793-3444
TELEFAX: (303) 793-3444
INFORMATION FOR SEQ ID NO: 219:
SEQUENCE CHARACTERISTICS:
LENGTH: 71 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

```

```

1  INFORMATION FOR SEQ ID NO: 5:
2  SEQUENCE CHARACTERISTICS:
3  LENGTH: 51 base pairs
4  TYPE: nucleic acid
5  STRANDEDNESS: single
6  topology: linear
7  MOLECULE TYPE: DNA
8  HYDROPHOBIC: NO
9  ANTI-SENSE: NO
10 FEATURE:
11 NAME/KEY: misc feature
12 LOCATION: 1..51
13 OTHER INFORMATION: cloning the linker peptide of Sequence ID No. 1."
14 US95/03742.5

Query Match: 88.4%; Score 12.4; DB %: Length 51;
Best Local Similarity: 92.9%; Prod. No. 2.8e+02;
Matches: 14; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;

QY 1 tetanotoxin 14
DB 14 TCAGGCTGGGATC 26

RESULT 6
US-09-341-829a-4_219/0
US-09-341-829a-4_219/0
Patent No. 5998142
GENERAL INFORMATION:
APPLICANT: GOLD et al.
TITLE OF INVENTION: SYSTEMATIC EVOLUTION OF LIBRARIES
BY EXPONENTIAL ENRICHMENT:
TITLE OF INVENTION: CHEM-SELEX
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratsch, L.L.C.
STREET: 8400 E. Front Street, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/341,829a-4_219/0
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 5998142
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,141
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/546,428
FILING DATE: 11-JUNE-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/117,991
FILING DATE: 9-SEPTEMBER-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/124,945
FILING DATE: 17-SEPTEMBER-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/199,597
FILING DATE: 22-FEBRUARY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/244,997
FILING DATE: 28-APRIL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/309,245
FILING DATE: 20-SEPTEMBER-1994
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 44,215
REFERENCE: 5998142
TELEPHONE: (303) 793-3444
TELEFAX: (303) 793-3444
INFORMATION FOR SEQ ID NO: 219:
SEQUENCE CHARACTERISTICS:
LENGTH: 71 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

```

1 MOLECULE TYPE: DNA
US-08-894-578-219

Query Match 88.6%; Score 12.4; DB 2; Length 71;
Best Local Similarity 92.9%; Pred. No. 2.9e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tetactctccgac 14
IIIIIIIIIIIIIIII

DB 22 TGTGCTCCGATC 9

RESULT 7

US-08-248-016-11/c
Sequence 11, Application US/09248016

Patent No. 5550109

GENERAL INFORMATION:

APPLICANT: Schonwetter, Barry S.

APPLICANT: Zasloff, Michael A.

TITLE OF INVENTION: Inducible Defensin Peptide From

Mammalian Epithelia

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

ADDRESSEE: Dunner

STREET: 1400 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/248,016

FILING DATE: 24 MAY-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Gaden, Stasia L.

REGISTRATION NUMBER: 36,228

REFERENCE/CHECKER NUMBER: 05387.0017-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4090

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 450 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-248-016-11

Query Match 88.6%; Score 12.4; DB 1; Length 350;
Best Local Similarity 92.9%; Pred. No. 2.9e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tetactctccgac 14
IIIIIIIIIIIIIIII

DB 258 TGTGCTCCGATC 245

RESULT 8

US-08-451-501-11/c

Sequence 11, Application US/08451501

Patent No. 5654738

GENERAL INFORMATION:

APPLICANT: Schonwetter, Barry S.

APPLICANT: Zasloff, Michael A.

TITLE OF INVENTION: Inducible Defensin Peptide From
Mammalian Epithelia

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

ADDRESSEE: Dunner

STREET: 1400 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/451,501

FILING DATE: 26-MAY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/248,016

FILING DATE: 24-MAY-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Fordis, Jean H.

REGISTRATION NUMBER: 30,907

REFERENCE/CHECKER NUMBER: 05387.3017-01000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 450 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-451-501-11

Query Match 88.6%; Score 12.4; DB 1; Length 450;
Best Local Similarity 92.9%; Pred. No. 2.9e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tetactctccgac 14
IIIIIIIIIIIIIIII

DB 258 TGTGCTCCGATC 245

RESULT 9

PCT-US95-06761-11/c

Sequence 11, Application PC/TUS9506761

GENERAL INFORMATION:

APPLICANT: Madadin Pharmaceuticals Inc.

APPLICANT: 5110 Campus Drive

APPLICANT: Plymouth Meeting, PA 19462

TITLE OF INVENTION: Inducible Defensin Peptide From

Mammalian Epithelia

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

ADDRESSEE: Dunner

STREET: 1400 I Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

1 CURRENT APPLICATION DATA:
 2 APPLICATION NUMBER: US 08/248,014
 3 FILING DATE: 24 MAY 1994
 4 PRIORITY APPLICATION DATA:
 5 APPLICATION NUMBER: US 08/248,014
 6 FILING DATE: 24 MAY 1994
 7 CLASSIFICATION:
 8 ATTORNEY/AGENT INFORMATION:
 9 NAME: Fabian, Gary R.
 10 REGISTRATION NUMBER: 44,876
 11 REFERENCE/INVENTOR INFORMATION:
 12 TELECOMMUNICATION INFORMATION:
 13 TELEPHONE: 402-408-4000
 14 TELEFAX: 402-408-4400
 15 INFORMATION FOR SEQ ID NO: 11:
 16 SEQUENCE CHARACTERISTICS:
 17 LENGTH: 350 base pairs
 18 TYPE: nucleic acid
 19 STRANDEDNESS: single
 20 TOPOLOGY: linear
 21 MOLECULE TYPE: DNA
 22 ORIGINAL SOURCE:
 23 INDIVIDUAL ISOLATE: clone D41-2
 24 US 08-611-757-48

1 Query Match: 88.6%; Score 12.4; DB 5; Length 350;
 2 Best Local Similarity: 92.9%; Freq. No. 2,900,023;
 3 Matches 14; Conservat 19; Mismatches 1; Indels 0; Gaps 0;
 4 QY 1 tetraoctacate 14
 5 DB 26A TETRAOCTACATE 245

1 RESULT 11
 2 US 08-611-757-48
 3 Sequence 48, Application US/08611757
 4 Patent No. 6,552,450
 5 GENERAL INFORMATION:
 6 APPLICANT: Kim, Junsuh P.
 7 APPLICANT: Boyes, Gregory R.
 8 APPLICANT: Waters, John
 9 APPLICANT: Clark, Rick; Clark, Yvonne
 10 APPLICANT: Young, Lawrence
 11 TITLE OF INVENTION: Reagents and Methods for
 12 TITLE OF INVENTION: Agents and Molecular Cloning Method
 13 NUMBER OF SEQUENCES: 106
 14 CORRESPONDENCE ADDRESS:
 15 ADDRESSEE: Behlinger & Associates
 16 STREET: 350 Cambridge Avenue, Suite 250
 17 CITY: Palo Alto
 18 STATE: CA
 19 COUNTRY: USA
 20 ZIP: 94306
 21 COMPUTER READABLE FORM:
 22 MEDIUM TYPE: floppy disk
 23 COMPUTER: IBM PC compatible
 24 OPERATING SYSTEM: PC DOS/MS DOS
 25 SOFTWARE: Patent to Release #1.9, Version #1.25
 26 CURRENT APPLICATION DATA:
 27 APPLICATION NUMBER: US/08/248,014
 28 FILING DATE:
 29 CLASSIFICATION:
 30 PRIOR APPLICATION DATA:
 31 APPLICATION NUMBER: 08/246,985
 32 FILING DATE: 20 MAY 1994
 33 ATTORNEY/AGENT INFORMATION:
 34 NAME: Fabian, Gary R.
 35 REGISTRATION NUMBER: 44,876
 36 REFERENCE/INVENTOR INFORMATION:
 37 TELECOMMUNICATION INFORMATION:
 38 TELEPHONE: (415) 324-0880
 39 TELEFAX: (415) 324-0960
 40 INFORMATION FOR SEQ ID NO: 48:
 41 SEQUENCE CHARACTERISTICS:
 42 LENGTH: 519 base pairs
 43 TYPE: nucleic acid
 44 STRANDEDNESS: double
 45 TOPOLOGY: linear
 46 MOLECULE TYPE: DNA
 47 HYDROPHETICAL: NO
 48 ORIGINAL SOURCE:
 49 INDIVIDUAL ISOLATE: clone D41-2
 50 US 08-611-757-48

1 Molecular Cloning Method
 2 TELECOMMUNICATION DATA:
 3 TELEPHONE: (415) 324-0880
 4 TELEFAX: (415) 324-0960
 5 INFORMATION FOR SEQ ID NO: 48:
 6 SEQUENCE CHARACTERISTICS:
 7 LENGTH: 519 base pairs
 8 TYPE: nucleic acid
 9 STRANDEDNESS: double
 10 TOPOLOGY: linear
 11 MOLECULE TYPE: DNA
 12 HYDROPHETICAL: NO
 13 ORIGINAL SOURCE:
 14 INDIVIDUAL ISOLATE: clone D41-2
 15 US 08-611-757-48

1 Query Match: 88.6%; Score 12.4; DB 2; Length 519;
 2 Best Local Similarity: 92.9%; Freq. No. 4,000,023;
 3 Matches 14; Conservat 19; Mismatches 1; Indels 0; Gaps 0;
 4 QY 1 tetraoctacate 14
 5 DB 394 TETRAOCTACATE 407

1 RESULT 11
 2 US 08-611-757-48
 3 Sequence 48, Application US/08611757
 4 Patent No. 6,552,450
 5 GENERAL INFORMATION:
 6 APPLICANT:
 7 APPLICANT:
 8 TITLE OF INVENTION: Reagents and Methods for
 9 TITLE OF INVENTION: Agents and Molecular Cloning Method
 10 NUMBER OF SEQUENCES: 106
 11 CORRESPONDENCE ADDRESS:
 12 ADDRESSEE: Behlinger & Associates
 13 STREET: 350 Cambridge Avenue, Suite 250
 14 CITY: Palo Alto
 15 STATE: CA
 16 COUNTRY: USA
 17 ZIP: 94306
 18 COMPUTER READABLE FORM:
 19 COMPUTER: IBM PC compatible
 20 OPERATING SYSTEM: PC DOS/MS DOS
 21 SOFTWARE: Patent to Release #1.9, Version #1.25
 22 CURRENT APPLICATION DATA:
 23 APPLICATION NUMBER: US/08/248,014
 24 FILING DATE:
 25 CLASSIFICATION:
 26 PRIOR APPLICATION DATA:
 27 APPLICATION NUMBER: US 08/246,985
 28 FILING DATE: 20 MAY 1994
 29 ATTORNEY/AGENT INFORMATION:
 30 NAME: Fabian, Gary R.
 31 REGISTRATION NUMBER: 44,876
 32 REFERENCE/INVENTOR INFORMATION:
 33 TELECOMMUNICATION INFORMATION:
 34 TELEPHONE: (415) 324-0880
 35 TELEFAX: (415) 324-0960
 36 INFORMATION FOR SEQ ID NO: 48:
 37 SEQUENCE CHARACTERISTICS:
 38 LENGTH: 519 base pairs
 39 TYPE: nucleic acid
 40 STRANDEDNESS: double
 41 TOPOLOGY: linear
 42 MOLECULE TYPE: DNA
 43 HYDROPHETICAL: NO
 44 ORIGINAL SOURCE:
 45 INDIVIDUAL ISOLATE: clone D41-2
 46 US 08-611-757-48

Query Match 88.6% Score 12.4; DB 5; Length 519;
 Best Local Similarity 92.9% Pred. No. 30-02;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 tetactctccatc 14
 DB 494 TGTGCTTGGCATC 407

RESULT 12
 US-09-446-026-327
 Sequence 12 Application US/08945026
 Patent No. 6034218
 GENERAL INFORMATION:
 APPLICANT: Reed, Steven G.
 APPLICANT: Dillon, David C.
 APPLICANT: Iwardzik, Daniel R.
 APPLICANT: Mitcham, Jonathan L.
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
 TITLE OF INVENTION: AND IMMUNODIAGNOSIS OF PROSTATE CANCER
 NUMBER OF SEQUENCES: 59
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED and BERRY LLP
 STREET: 6400 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104-7092

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.40
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/02/946,026
 FILING DATE: 07-OCT-1997
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Makl, David J.
 REGISTRATION NUMBER: 41,492
 REFERENCE/INVENT NUMBER: 210121.424C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 768 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-946-026-32
 Query Match 88.6% Score 12.4; DB 4; Length 768;
 Best Local Similarity 92.9% Pred. No. 30-02;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 tetactctccatc 14
 DB 62 TGTGCTTGGCATC 49

RESULT 13
 US-09-475-316A-14
 Sequence 14 Application US/09475016A
 Patent No. 6210942
 GENERAL INFORMATION:
 APPLICANT: Lewis, No. 6210942man G.
 APPLICANT: Davin, Laurence B.
 APPLICANT: Binkova-Fostova, Alena T.
 APPLICANT: Fujita, Masayuki
 APPLICANT: Ganq, David R.

APPLICANT: Sarkanen, Simo
 APPLICANT: Ford, Joshua D.
 TITLE OF INVENTION: RECOMBINANT FIBROBLAST/LARGICININ-G RELAXTASES,
 TITLE OF INVENTION: RECOMBINANT DIKISIN PROTEINS AND METHODS OF USE
 FILE REFERENCE: WSUR-1-14793
 CURRENT APPLICATION NUMBER: US/09/475,316A
 CURRENT FILING DATE: 1999-12-30
 PRIOR APPLICATION NUMBER: 09/407,653
 PRIOR FILING DATE: 1999-05-07
 PRIOR APPLICATION NUMBER: PCT/0097/20391
 PRIOR FILING DATE: 1997-11-07
 PRIOR APPLICATION NUMBER: 60/054,380
 PRIOR FILING DATE: 1997-07-31
 PRIOR APPLICATION NUMBER: 60/030,522
 PRIOR FILING DATE: 1996-11-08
 NUMBER OF SEQ ID NOS: 122
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 14
 LENGTH: 858
 TYPE: DNA
 ORGANISM: Forsythia x intermedia
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (19)..(574)
 US-09-475-316A-14

Query Match 88.6% Score 12.4; DB 4; Length 858;
 Best Local Similarity 92.9% Pred. No. 40-02;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 tetactctccatc 14
 DB 53 tetactctccatc 66

RESULT 14
 US-08-910-731-1
 Sequence 14 Application US/08910731
 Patent No. 5942440
 GENERAL INFORMATION:
 APPLICANT: CHATKIEB, DEB K.
 APPLICANT: SHANDILYA, HARINI
 TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
 STREET: 1100 NEW YORK AVE., N.W., SUITE 600
 CITY: WASHINGTON
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/910,731
 FILING DATE: (Herewith)
 CLASSIFICATION: 445
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/795,395
 FILING DATE: 04-FEB-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/794,546
 FILING DATE: 03-FEB-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/024,057
 FILING DATE: 16-AUG-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: RSMOND, ROBERT W.
 REGISTRATION NUMBER: 32,894

1 REFERENCE/DESCRIPT NUMBER: 0942, 4440003
 2 TELECOMMUNICATION INFORMATION:
 3 TELEPHONE: 202-471-2500
 4 TELEFAX: 202-471-2540
 5 INFORMATION FOR SEQ ID NO: 1:
 6 SEQUENCE CHARACTERISTICS:
 7 LENGTH: 1471 base pairs
 8 TYPE: nucleic acid
 9 STRANDEDNESS: both
 10 TOPOLGY: both
 11 MOLECULE TYPE: cDNA
 12 FEATURE:
 13 NAME/KEY: CDS
 14 LOCATION: 1..1408
 15 US 08 910 741 1

Query Match: 88.6%; Score 12.4; DB Z: Length 1471;
 Best Local Similarity: 92.9%; Prod. No. 40002;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tetraocteacate 14
 DB 167 tetraocteacate 180

RESULT 15

US 08 910 741 7
 2 Sequence 7, Application US/08910741
 3 Patent No. 5932446

GENERAL INFORMATION

1 APPLICANT: CHATTERJEE, DEB K.

2 APPLICANT: SHANILYA, HARINI

3 TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof

4 NUMBER OF SEQUENCES: 15

5 CORRESPONDENCE ADDRESS:

6 ADDRESSEE: STEPHEN KESSLER, GOLDSSTEIN & FOX P.L.L.C.

7 STREET: 1100 NEW YORK AVE., N.W., SUITE 600

8 CITY: WASHINGTON

9 STATE: D.C.

10 COUNTRY: USA

11 ZIP: 20005, 5944

12 COMPUTER READABLE FORM:

13 MEDIA TYPE: Floppy disk

14 COMPUTER: IBM PC compatible

15 OPERATING SYSTEM: PC DOS/MS-DOS

16 SOFTWARE: Patented Release #1.0, Version #1.40

17 CURRENT APPLICATION DATA:

18 APPLICATION NUMBER: 08/910,741

19 FILING DATE: (Berwick)

20 CLASSIFICATION: 435

21 PRIOR APPLICATION DATA:

22 APPLICATION NUMBER: 08/795,495

23 FILING DATE: 04-FEB-1997

24 PRIOR APPLICATION DATA:

25 APPLICATION NUMBER: 08/794,546

26 FILING DATE: 04-FEB-1997

27 PRIOR APPLICATION DATA:

28 APPLICATION NUMBER: 60/024,057

29 FILING DATE: 16-AUG-1996

30 ATTORNEY/AGENT INFORMATION:

31 NAME: ESMOND, ROBERT W.

32 REGISTRATION NUMBER: 42,893

33 REFERENCE/SEQ ID NUMBER: 0942, 4440003

34 TELECOMMUNICATION INFORMATION:

35 TELEPHONE: 202-471-2500

36 TELEFAX: 202-471-2540

37 INFORMATION FOR SEQ ID NO: 7:

38 SEQUENCE CHARACTERISTICS:

39 LENGTH: 1471 base pairs

40 TYPE: nucleic acid

41 STRANDEDNESS: both

42 TOPOLOGY: both

1 MOLECULE TYPE: cDNA
 US 08 910 741 7

Query Match: 88.6%; Score 12.4; DB Z: Length 1471;
 Best Local Similarity: 92.9%; Prod. No. 40002;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tetraocteacate 14
 DB 167 tetraocteacate 180

Search completed: May 1, 2001, 03:54:26
 Job time: 7276 sec

Copyright (c) 1993 - 2000 CompuLink Ltd.

Software version 4.5

CM nucleotide nucleotide search, using sw model

Run on: May 1, 2001, 03:04:03, Search time 4766.75 seconds

(without alignments)

42,299 Million cell updates/sec

Title: US 09-341-829A-4 copy_1_14

Posttest score: 14

Sequence: 1 retgctctcagcgc 14

Scoring table: IDENTIFY NUC

Gapop 10,0 / Gapext 1,0

Searches: 92,299 17 steps, 46,904,929 matches

Total number of hits satisfying chosen parameters: 19247044

Minimum hit seq length: 0

Maximum hit seq length: 200000000

Post processing: Minimum Match 0%

Maximum Match 100%

Fastest First 4% summaries

Database:

EST:*

1: qb_ost11:*

2: qb_ost2:*

3: qb_ost3:*

4: qb_ost4:*

5: qb_ost5:*

6: qb_ost6:*

7: qb_ost7:*

8: qb_ost8:*

9: qb_ost9:*

10: qb_ost10:*

11: qb_ost11:*

12: qb_ost12:*

13: qb_ost13:*

14: qb_ost14:*

15: qb_ost15:*

16: qb_ost16:*

17: qb_ost17:*

18: qb_ost18:*

19: qb_ost19:*

20: qb_ost20:*

21: qb_ost21:*

22: qb_ost22:*

23: qb_ost23:*

24: qb_ost24:*

25: qb_ost25:*

26: qb_ost26:*

27: qb_ost27:*

28: qb_ost28:*

29: qb_ost29:*

30: qb_ost30:*

31: qb_ost31:*

32: qb_ost32:*

33: qb_ost33:*

34: qb_ost34:*

35: qb_ost35:*

36: qb_ost36:*

37: qb_ost37:*

38: qb_ost38:*

39: qb_ost39:*

40: qb_ost40:*

41: qb_ost41:*

42: qb_ost42:*

43: qb_ost43:*

44: qb_ost44:*

45: qb_ost45:*

46: qb_ost46:*

47: qb_ost47:*

44: cm_osthum10:*

45: cm_osthum11:*

46: cm_osthum12:*

47: cm_osthum13:*

48: cm_osthum14:*

49: cm_osthum15:*

50: cm_osthum16:*

51: cm_osthum17:*

52: cm_osthum18:*

53: cm_osthum19:*

54: cm_osthum20:*

55: cm_osthum21:*

56: cm_osthum22:*

57: cm_osthum23:*

58: cm_osthum24:*

59: cm_osthum25:*

60: cm_osthum26:*

61: cm_osthum27:*

62: cm_osthum28:*

63: cm_osthum29:*

64: cm_osthum30:*

65: cm_osthum31:*

66: cm_osthum32:*

67: cm_osthum33:*

68: cm_osthum34:*

69: cm_osthum35:*

70: cm_osthum36:*

71: cm_osthum37:*

72: cm_osthum38:*

73: cm_osthum39:*

74: cm_osthum40:*

75: cm_osthum41:*

76: cm_osthum42:*

77: cm_osthum43:*

78: cm_osthum44:*

79: cm_osthum45:*

80: cm_osthum46:*

81: cm_osthum47:*

82: cm_osthum48:*

83: cm_osthum49:*

84: cm_osthum50:*

85: cm_osthum51:*

86: cm_osthum52:*

87: cm_osthum53:*

88: cm_osthum54:*

89: cm_osthum55:*

90: cm_osthum56:*

91: cm_osthum57:*

92: cm_osthum58:*

93: cm_osthum59:*

94: cm_osthum60:*

95: cm_osthum61:*

96: cm_osthum62:*

97: cm_osthum63:*

98: cm_osthum64:*

99: cm_osthum65:*

100: cm_osthum66:*

101: cm_osthum67:*

102: qb_ost25:*

103: qb_ost26:*

104: qb_ost27:*

105: qb_ost28:*

106: qb_ost29:*

107: qb_ost30:*

108: qb_ost31:*

109: qb_ost32:*

110: qb_ost33:*

111: qb_ost34:*

112: qb_ost35:*

113: qb_ost36:*

114: qb_ost37:*

115: qb_ost38:*

116: qb_ost39:*

117: qb_est48:*
 118: qb_est49:*
 119: qb_est50:*
 120: qb_est51:*
 121: qb_est52:*
 122: qb_est53:*
 123: qb_est54:*
 124: qb_est55:*
 125: qb_est56:*
 126: qb_est57:*
 127: qb_est58:*
 128: qb_est59:*
 129: qb_est60:*
 130: qb_est61:*
 131: qb_est62:*
 132: qb_est63:*
 133: qb_est64:*
 134: qb_est65:*
 135: qb_est66:*
 136: qb_est67:*
 137: qb_est68:*
 138: qb_est69:*
 139: qb_est70:*
 140: qb_est71:*
 141: qb_est72:*
 142: qb_est73:*
 143: qb_est74:*
 144: qb_est75:*
 145: qb_est76:*
 146: qb_est77:*
 147: qb_est78:*
 148: qb_est79:*
 149: qb_est80:*
 150: qb_est81:*
 151: qb_est82:*
 152: qb_est83:*
 153: qb_est84:*
 154: qb_est85:*
 155: qb_est86:*
 156: qb_est87:*
 157: qb_est88:*
 158: qb_est89:*
 159: qb_est90:*
 160: qb_est91:*
 161: qb_est92:*
 162: qb_est93:*
 163: qb_est94:*
 164: qb_est95:*
 165: qb_est96:*
 166: qb_est97:*
 167: qb_est98:*
 168: qb_est99:*
 169: qb_est100:*
 170: qb_est101:*
 171: qb_est102:*
 172: qb_est103:*
 173: qb_est104:*
 174: qb_est105:*
 175: qb_est106:*
 176: qb_est107:*
 177: qb_est108:*
 178: qb_est109:*
 179: qb_est110:*
 180: qb_est111:*
 181: qb_est112:*
 182: qb_est113:*
 183: qb_est114:*
 184: qb_est115:*
 185: qb_est116:*
 186: qb_est117:*
 187: qb_est118:*
 188: qb_est119:*
 189: qb_est120:*

190: em_qss_pln1:*
 191: em_qss_pln2:*
 192: em_qss_pro:*
 193: em_qss_rod1:*
 194: em_qss_rod2:*
 195: em_qss_rod3:*
 196: em_qss_rod4:*
 197: em_qss_rod5:*
 198: em_qss_vrt1:*
 199: em_qss_vrt2:*
 200: em_qss_vrt3:*
 201: qb_qss1:*
 202: qb_qss2:*
 203: qb_qss3:*
 204: qb_qss4:*
 205: qb_qss5:*
 206: qb_qss6:*
 207: qb_qss7:*
 208: qb_qss8:*
 209: qb_qss9:*
 210: qb_qss10:*
 211: qb_qss11:*
 212: qb_qss12:*
 213: qb_qss13:*
 214: qb_qss14:*
 215: qb_qss15:*
 216: qb_qss16:*
 217: qb_qss17:*
 218: qb_qss18:*
 219: qb_qss19:*
 220: qb_qss20:*
 221: qb_qss21:*
 222: qb_qss22:*
 223: qb_qss23:*
 224: qb_qss24:*
 225: qb_qss25:*
 226: qb_qss26:*
 227: qb_qss27:*
 228: qb_qss28:*
 229: qb_qss29:*
 230: qb_qss30:*
 231: qb_qss31:*
 232: qb_qss32:*
 233: qb_qss33:*
 234: qb_qss34:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | Query | Match | Length | DB | ID | Description |
|--------|-----|-------|-------|-------|--------|-----------|-----------|-------------|
| | 1 | 14 | 100.0 | 168 | 117 | AW579092 | AW579092 | RCO_C1046 |
| | 2 | 14 | 100.0 | 292 | 124 | B0055977 | B0055977 | B0055977 |
| | 3 | 14 | 100.0 | 402 | 106 | A0040252 | A0040252 | A0040252 |
| | 4 | 14 | 100.0 | 429 | 104 | A1282204 | A1282204 | AA3A-AA0- |
| | 5 | 14 | 100.0 | 472 | 3 | AA155178 | AA155178 | mt_R1107.1 |
| | 6 | 14 | 100.0 | 478 | 165 | HE346096 | HE346096 | sf19d06.y |
| | 7 | 14 | 100.0 | 525 | 166 | HE367560 | HE367560 | P11.9_F03 |
| | 8 | 14 | 100.0 | 541 | 215 | A2068067 | A2068067 | RFC1_23-3 |
| | 9 | 14 | 100.0 | 545 | 174 | RC084012 | RC084012 | RC084012 |
| | 10 | 14 | 100.0 | 548 | 2 | AA141767 | AA141767 | CR02398.c |
| | 11 | 14 | 100.0 | 554 | 239 | AZ221492 | AZ221492 | IM0342103 |
| | 12 | 14 | 100.0 | 570 | 172 | B0071181 | B0071181 | B0071181 |
| | 13 | 14 | 100.0 | 597 | 18 | A1294860 | A1294860 | LP08330.5 |
| | 14 | 14 | 100.0 | 616 | 225 | AZ665079 | AZ665079 | 1R3-5R17 |
| | 15 | 14 | 100.0 | 618 | 232 | FR0041522 | FR0041522 | AL129014 |
| | 16 | 14 | 100.0 | 619 | 242 | FR0042367 | FR0042367 | Fuq1_r0b1 |
| | 17 | 14 | 100.0 | 623 | 22 | A1601278 | A1601278 | Fuq1_r0b1 |
| | 18 | 14 | 100.0 | 638 | 117 | AW611303 | AW611303 | 4599405.X |
| | 19 | 14 | 100.0 | | | | | AW611303 |

/clone_lib "wipfn full-length enriched, 12 days embryo
 male wollian duct"
 /sex "male"
 /issue_type "wollian duct includes surrounding region"
 /dev_stage "12 days embryo"
 /lab_host "bhl08"
 /core "site_1; SalI; Site 2; BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Wistar
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN, Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGAGAGATGCAAGAGCTGTTTTTTTTTTTTTTT 3'], cDNA was
 prepared by using probalase thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. Second strand cDNA was prepared with the
 primer adaptor of sequence [5'
 GAGAGAGATGTCGACTTAATTAATTAATGCGCGCGCGCC 3']. cDNA
 was cloned into the XbaI and BamHI sites. Vector: a
 modified pluescript KS(+) after bulk excision from lambda
 phage 1. Cloning sites, 5' end: SalI; 3' end: BamHI"
 BASE COUNT 63 a 68 c 56 g 105 t
 ORIGIN
 1 tetgenctccgacatc 14
 |||||||
 db 43 tctgctccgacatc 46

Query Match 100.0%; Score 14; DB 124; Length 292.
 Best Local Similarity 100.0%; Pred. No. 2.5e+03;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 tetgenctccgacatc 14
 |||||||
 DB 43 tctgctccgacatc 46
 RESULT 3
 LOCUS A0040252/c
 DEFINITION A0040252 Mouse four-cell-embryo cDNA Mus musculus cDNA clone
 J0802503 3', mRNA sequence.
 ACCESSION A0040252
 VERSION A0040252.1 GI:3954072
 KEYWORDS EST
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 KO: M.S.H.; Kitchen, J.K.; Wang, X.; Threlk, J.A.; Sun, T.; DePalma, C.E.
 Liang, Y.; Kuruk, G.J.; Sharara, R.; Lim, M.K. and Doi H
 Systematic analyses of genes expressed in 4-cell mouse embryo (The
 EPAT/Doi Project at Wayne State University)
 CONTACT: Hirotsumi Doi
 Doi Biosymmetry Project, EPAT
 Japan Science and Technology Corporation (JST)
 Wako Marine East 12F, 3-6 Nakase, Mihama-ku, Chiba 261-71, Japan
 Email: hdoibio@jst.go.jp.
 Location/Qualifiers
 1..402
 /organism "Mus musculus"
 /strain "w57u/6J"
 /db_xref="taxon:10090"
 /clone="J0802503"
 /clone_lib="mouse four-cell-embryo cDNA"
 /dev_stage="four-cell-embryo"
 141 a 76 c 64 g 120 t 1 others
 BASE COUNT
 ORIGIN
 Query Match 100.0%; Score 14; DB 136; Length 402.
 Best Local Similarity 100.0%; Pred. No. 2.5e+03;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 1 tctgenctccgacatc 14
 |||||||
 DB 243 tctgctccgacatc 220
 RESULT 4
 LOCUS A2282204/c
 DEFINITION 43A AAG-B 05 R Anopheles gambiae immune competent 43A Anophelies
 gambiae cDNA clone 43A AAG-B 05, mRNA sequence.
 ACCESSION A2282204
 VERSION A2282204.1 GI:6930083
 KEYWORDS EST
 SOURCE African malaria mosquito.
 ORGANISM Anopheles gambiae
 Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta; Diptera;
 Pterygota; Neoptera; Endopterygota. Diptera, Nematocera; Culicoidae
 ; Anopheles.
 1 (bases 1 to 439)
 Dimopoulos, G.; Casavant, T.L.; Chang, S.; Schaefer, J.; Roberts, J.;
 Donohue, M.; Schmitz, J.; Hones, V.; Bork, P.; Ansorge, W.; Soares, M.B.
 and Kafatos, F.C.
 Anopheles gambiae pilot gene discovery project: identification of
 mosquito innate immunity genes from expressed sequence tags
 generated from immune-competent cell lines
 Proc. Natl. Acad. Sci. U.S.A. 97 (14), 9619-9624 (2000)
 20300950
 CONTACT: Dimopoulos G
 Follis G. Kafatos Laboratory
 European Molecular Biology Laboratory
 Meyerhofstrasse 1, 69117 Heidelberg, Germany.
 Location/Qualifiers
 1..439
 /organism "Anopheles gambiae"
 /strain="4A 1/1"
 /db_xref="taxon:7165"
 /clone "4A3A-AAG-B-09"
 /clone_lib "Anopheles gambiae immune competent 4A3A"
 /cell_line="immune competent 4A3A"
 /lab_host="E. coli DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: EcoRI; Site 2: NotI; sequenced from
 forward priming site which reads from the 3' end of the
 cDNA. The 4A3A is a directionally cloned and normalized
 cDNA library that was constructed from the 4A3A cell line
 oligo-1 primed cDNA according to: Ronaldo, Louisa & Soares
 (1996) : Normalization and Subtraction: Two approaches to
 Facilitate Gene Discovery, Genome Research 6, 791-800."
 BASE COUNT 144 a 112 c 122 g 61 t
 ORIGIN
 Query Match 100.0%; Score 14; DB 104; Length 449;
 Best Local Similarity 100.0%; Pred. No. 2.5e+04;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 tctgenctccgacatc 14
 |||||||
 DB 483 tctgctccgacatc 370
 RESULT 5
 LOCUS AAI55178/c
 DEFINITION T81197.1, StrataGene mouse heart (497376) Rat musculus (RAT clone)
 IMAGE:603877 5' similar to gb:03272 FIBRILLIN 2 (HUMAN); gb:022494
 Mus musculus fibrillin-1 (MOUSE); mRNA sequence.
 ACCESSION AAI55178
 VERSION AAI55178.1 GI:1726802
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 167 a 100 c 111 q 176 t
ORIGIN
Query Match 100.0%; Score 14; DB 219; Length 554;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tetgctccgcacate 14
|||||
DB 496 TGTGCTCCGCATC 508

RESULT 12
LOCUS B071181/c 570 bp mRNA EST 26-JAN-2001
DEFINITION H3095F04 3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
ACCESSION B071181 3', mRNA sequence,
VERSION B071181.1 GI:12754750
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
1 (bases 1 to 570)
REFERENCE Kargul,G.J., Indekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
J.L.S., Carter,M.G. and Ko,M.S.H.
Verification and initial annotation of NIA mouse 15K cDNA clone set
Unpublished (2001)
OTHER ESTs: H3095F04-5
CONTACT: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
345 Cassell Drive, Suite 4000, Baltimore, MD 21224-6920, USA
Email: edna@lqsm-grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://lqsm-grc.nia.nih.gov/ncg/15k.html> for details.
Plato: H3095 row: F column: 04
Seq primer: -21M13 Forward
High quality sequence stop: 570
POLYA Yes.

FEATURES
source
1..570 Location/Qualifiers
/organism="Mus musculus"
/strain="acc70/6.1"
/db_xref="taxon:10090"
/clone="H3095F04"
/cclone_lib="NIA Mouse 15K cDNA clone set"
/sex="clones arrayed from a mixture of cDNA libraries"
/cov_stage="clones arrayed from a mixture of cDNA libraries"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearrayed set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryo, extraembryonic part of E7.5 embryo, and E12.5 female mesonephros/ovary) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U.S.A. 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1747-1749; (3) Genome-wide mapping of

unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in blast complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

BASE COUNT 188 a 122 c 92 q 168 t
ORIGIN
Query Match 100.0%; Score 14; DB 172; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tetgctccgcacate 14
|||||
DB 233 TGTGCTCCGCATC 220

RESULT 13
LOCUS A1294860 597 bp mRNA EST 01-DEC-1998
DEFINITION LP08330-5prime 1P Drosophila melanogaster larval-early pupal polyA
ACCESSION A1294860 Drosophila melanogaster cDNA clone LP08330 5prime, mRNA sequence.
VERSION A1294860.1 GI:3944267
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila;
1 (bases 1 to 597)
REFERENCE Harvey,D., Honan,L., Evans-Bulch,M., Pendleton,J., Su,D., Blockstein
P., Lewis,S. and Rubin,G.M.
Black/UMI Drosophila EST Project
Unpublished (1997)
CONTACT: Harvey, D.
G. M. Rubin-Molecular and Cell Biology
University of California Berkeley
539 LSA, Berkeley, CA 94720-3200, USA
Fax: 510 643 9947
Email: http://www.fruitfly.org/est_estat/fly_berkeley7.edu
Plato: 84 row: C column: 6
High quality sequence stop: 443.
Location/Qualifiers
1..597
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="LP08330"
/cclone_lib="1P Drosophila melanogaster larval-early pupal polyA"
/sex="male and female"
/cov_stage="larvae-pupae"
/lab_host="DH5-alpha"
/note="Organ whole body; Vector: pOT2; Site_1: EcoRI;
Site_2: XbaI; Sized fractionated cDNAs were directly ligated into pOT2. Plasmid cDNA library."

FEATURES
source
1..597 Location/Qualifiers
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="LP08330"
/cclone_lib="1P Drosophila melanogaster larval-early pupal polyA"
/sex="male and female"
/cov_stage="larvae-pupae"
/lab_host="DH5-alpha"
/note="Organ whole body; Vector: pOT2; Site_1: EcoRI;
Site_2: XbaI; Sized fractionated cDNAs were directly ligated into pOT2. Plasmid cDNA library."
BASE COUNT 133 a 143 c 181 q 140 t
ORIGIN
Query Match 100.0%; Score 14; DB 18; Length 597;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tetgctccgcacate 14
|||||
DB 334 TGTGCTCCGCATC 347

RESULT 14
LOCUS AZ465079/c 616 bp DNA GSS 14 DEC-2000

GenCore version 4.1b
Copyright (c) 1993-2000 CompuLink Ltd.

RM nucleotide nucleotide search, using sw model

Run on: May 1, 2001, 09:03:52 : Search time 25:12:48 Seconds
(without alignments)

131 loci US-09-341-829a-4_copy_250_755
Sequences: 506
1 nucleotide query sequence.....created at c:\cqa\cqa 506

Search table: IDENTITY NUC
Gapop 10.0 : Gapex 1.0

Searches: 1283235 seqs, 737892952 residues 2566470
Total number of hits satisfying chosen parameters:

Minimum hit seq length: 6
Maximum hit seq length: 200000000

Post processing: Minimum Match: 0%
Fastest First: 45 summaries

Database: 1: db_ba1: *
2: db_ba2: *
3: db_ba3: *
4: db_ba4: *
5: db_ba5: *
6: db_ba6: *
7: db_ba7: *
8: db_ba8: *
9: db_ba9: *
10: db_ba10: *
11: db_ba11: *
12: db_ba12: *
13: db_ba13: *
14: db_ba14: *
15: db_ba15: *
16: db_ba16: *
17: db_ba17: *
18: db_ba18: *
19: db_ba19: *
20: db_ba20: *
21: db_ba21: *
22: db_ba22: *
23: db_ba23: *
24: db_ba24: *
25: db_ba25: *
26: db_ba26: *
27: db_ba27: *
28: db_ba28: *
29: db_ba29: *
30: db_ba30: *
31: db_ba31: *
32: db_ba32: *
33: db_ba33: *
34: db_ba34: *
35: db_ba35: *
36: db_ba36: *
37: db_ba37: *
38: db_ba38: *
39: db_ba39: *
40: db_ba40: *
41: db_ba41: *
42: db_ba42: *
43: db_ba43: *

44: cm_ov: *
45: cm_ov: *
46: cm_ov: *
47: cm_ov: *
48: cm_ov: *
49: cm_ov: *
50: cm_ov: *
51: cm_ov: *
52: cm_ov: *
53: db_ba1: *
54: db_ba2: *
55: db_ba3: *
56: db_ba4: *
57: db_ba5: *
58: db_ba6: *
59: db_ba7: *
60: db_ba8: *
61: db_ba9: *
62: db_ba10: *
63: db_ba11: *
64: db_ba12: *
65: db_ba13: *
66: db_ba14: *
67: db_ba15: *
68: db_ba16: *
69: db_ba17: *
70: db_ba18: *
71: db_ba19: *
72: db_ba20: *
73: db_ba21: *
74: db_ba22: *
75: db_ba23: *
76: db_ba24: *
77: db_ba25: *
78: db_ba26: *
79: db_ba27: *
80: db_ba28: *
81: db_ba29: *
82: db_ba30: *
83: db_ba31: *
84: db_ba32: *
85: db_ba33: *
86: db_ba34: *
87: db_ba35: *
88: db_ba36: *
89: db_ba37: *
90: db_ba38: *
91: db_ba39: *
92: db_ba40: *
93: db_ba41: *
94: db_ba42: *
95: db_ba43: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|-----------|-------------|
| 1 | 506 | 100.0 | 993 | 9 | AX024689 | Sequence |
| 2 | 506 | 100.0 | 993 | 92 | BSA012835 | Homo sapi |
| 3 | 506 | 100.0 | 1002 | 9 | AB425335 | Sequence |
| 4 | 506 | 100.0 | 1002 | 92 | BSA223040 | Homo sapi |
| 5 | 422 | 83.4 | 112874 | 78 | AF277315 | Homo sapi |
| 6 | 418.8 | 82.8 | 9245 | 91 | BS223093 | Homo sapi |
| 7 | 365.6 | 72.3 | 873 | 92 | BSA275978 | Homo sapi |
| 8 | 305.2 | 60.3 | 2630 | 92 | BSA275977 | Homo sapi |
| 9 | 305.2 | 60.3 | 112874 | 78 | AF277315 | Homo sapi |
| 10 | 219 | 43.3 | 679 | 9 | AX024685 | Sequence |
| 11 | 219 | 43.3 | 679 | 92 | BSA012833 | Homo sapi |

| | | | | | | |
|----|------|------|--------|----|-------------|------------|
| 12 | 219 | 43.3 | 755 | 9 | AR042536 | Sequence |
| 13 | 219 | 43.3 | 755 | 92 | AJ224041 | Homo sapi |
| 14 | 219 | 43.3 | 767 | 9 | HS024687 | Sequence |
| 15 | 219 | 43.3 | 767 | 92 | HS013834 | Homo sapi |
| 16 | 192 | 48.1 | 237 | 9 | AK024532 | Sequence |
| 17 | 155 | 40.6 | 752 | 9 | AR037873 | Sequence |
| 18 | 155 | 40.6 | 752 | 9 | AX024763 | Sequence |
| 19 | 155 | 40.6 | 752 | 9 | AX024593 | Sequence |
| 20 | 155 | 40.6 | 752 | 93 | HS087459 | Human auto |
| 21 | 155 | 40.6 | 755 | 9 | AR042537 | Sequence |
| 22 | 155 | 40.6 | 755 | 92 | HS003149 | Homo sapi |
| 23 | 155 | 40.6 | 806 | 98 | AF038567 | Homo sapi |
| 24 | 68.4 | 13.5 | 21947 | 93 | H0MFLAG6620 | Homo sapi |
| 25 | 45.6 | 9.2 | 599 | 92 | HS11182 | Homo sapi |
| 26 | 47.8 | 9.7 | 145918 | 79 | AL162590 | Homo sapi |
| 27 | 49.6 | 7.8 | 34478 | 93 | HIMULAC19 | Human cos |
| 28 | 89.6 | 7.8 | 158414 | 86 | AC005940 | Homo sapi |
| 29 | 89.6 | 7.8 | 180494 | 62 | AC011189 | Homo sapi |
| 30 | 89.6 | 7.8 | 191674 | 68 | AC022875 | Homo sapi |
| 31 | 89.4 | 7.8 | 927 | 53 | CNS006XJF | End of |
| 32 | 89 | 7.7 | 143075 | 73 | AF205525 | Homo sapi |
| 33 | 88.8 | 7.7 | 181047 | 65 | AC016765 | Homo sapi |
| 34 | 88.2 | 7.5 | 96964 | 97 | AC031664 | Homo sapi |
| 35 | 88.2 | 7.5 | 39915 | 87 | AC010512 | Homo sapi |
| 36 | 88.2 | 7.5 | 203491 | 86 | AC007156 | Homo sapi |
| 37 | 88.2 | 7.5 | 216631 | 95 | AC005332 | Homo sapi |
| 38 | 87.8 | 7.5 | 169611 | 78 | AC087891 | Mus muscu |
| 39 | 87.4 | 7.4 | 167860 | 81 | AL490916 | Homo sapi |
| 40 | 87.2 | 7.4 | 43966 | 73 | AC034180 | Homo sapi |
| 41 | 87.2 | 7.4 | 15717 | 61 | AC010336 | Homo sapi |
| 42 | 87.2 | 7.1 | 170945 | 78 | AL137789 | Homo sapi |
| 43 | 87.2 | 7.4 | 198677 | 1 | AF001803 | Homo sapi |
| 44 | 87.2 | 7.4 | 345116 | 2 | AP003003 | Delinorace |
| 45 | 87 | 7.3 | 188838 | 65 | AC016877 | Homo sapi |

ALIGNMENT

```

RESEQ_1
AX024689
16095
DEFINITION
AX024689
ACCESSION
AX024689.1
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 993)
Aarnoudse, C.A. and Schrier, P.
Camel, an alternative translation product of the tumour antigen
tag 1
PATENT: EP 1001022 A 5 17 MAY 2009;
BOEHRINGER INGELHEIM INT (DE) ; UNIV HOSPITAL LEIDEN (NL)
FEATURES
Source
5'UTR
CDS
1943..2953
3443..1831

```


| | |
|--|---|
| ORGANISM | Homo sapiens |
| REFERENCE | Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. |
| AUTHORS | Lethé B.G. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (06 JAN 1998) Lethé B.G., Brussels Branch, Ludwig Institute for Cancer Research, 74, avenue Hippocrate, B-1200 Bruxelles, BELGIUM |
| REFERENCE | 2 (bases 1 to 1002) |
| AUTHORS | Lethé B., Lucas S., Michaux J., De Smet C., Godelaine D., Serrano A., De Plaen F. and Boon T. |
| TITLE | LAGE-1, a new gene with tumor specificity |
| JOURNAL | Int. J. Cancer 76 (6), 903-908 (1998) |
| MEDLINE | 98289662 |
| FEATURES | Location/Qualifiers 1..1002 organism "Homo sapiens" /isolate "individual PB373" /db_xref "Gene:494" /chromosome "X" /cell_line "LB373-MEL" /cell_type "melanoma" /map_q28 /feature "L1-1" /feature "Lage-1" 65..976 /gene "LAGE-1" 65..997 /gene="LAGE-1" /region_start 1 /product "LAGE-1 protein" /protein_id "CAAL1049.1" /db_xref "UniProt:P55593" /translation "MGAAVGGGSGSTFATSPSSQITFGPGNAKIFAPAFPEPFA PRGAARANGDPGRAGDHPWCAACAWSCQCAPPNSPIQLHTMPDSSEPMIA ELVRLLSPDAHLPPGCAYLKDTVTSCNLPMSVPDGRCACRMLVVGVWGLGSASP EGAKAPLPTDKRKVSFPWPSFTGPDPPEAGGDCRCROVAENMFSAPII" |
| polyA_signal | 971..976 |
| polyA_site | 993 |
| BASE COUNT | 179 a 302 c 336 g 185 t |
| ORIGIN | |
| Query Match | 100.0%; Score 506; DB 92; Length 1002; |
| Best Local Similarity | 100.0%; Pred. No. 2,1e+104; |
| Matches 506; conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| QQ | 1 aaatcccatcattaccctactgcggccatatgaataccctacqcccccaaaag 60 |
| bb | |
| 250 | GAGTCGCGATGGGCTGTGTTTCTGTGTAATGGGTTGGTGGGCGCGAG 309 |
| QQ | 61 qcccdacagcccctgcttcattgcacatcacatgtcttctcgtgcccatgaaaac 120 |
| bb | |
| 310 | GTGCGAAGAGCCCTTGCTTAGTTGCATCACCATCCTCTCTCTCGCGCATGCAACC 369 |
| QQ | 121 qaagctatcccgaaagatctqtccccgaalgcgcgcacctccccccagagaggaatt 180 |
| bb | |
| 370 | GGAGCTGTCTCCGCAAGATCTGTGCCGGATTGGCGAATGGCGAGATGCGGCGGT 429 |
| QQ | 181 tcttgaaagctccctgtatcaggcaactcctattatgctgctttggaacaaagaaag 240 |
| bb | |
| 440 | TCTCAAAGACTATCCCTGTGGGCTAACGTATCTTTAATGAGTCTGACATCAAG 486 |
| QQ | 241 qaagagctdqqccatqaagtgttgaatgatggagctccccccagagaggaaga 300 |
| bb | |
| 490 | GGAAAGGCTGGAGATGAGATGGATGGATGGATGGATGGATGGATGGATGGATGGATGG 549 |
| QQ | 301 gaagagagcttatgattctgagatccacagctgaatttcagatataagagatgplac 360 |
| bb | |
| 550 | GCAGAAAGCTATGATCTCAGAAAGCCAAAACAAGGTTTCAGAACAGAGCCCTTAC 609 |

DR WPI: 1998-427451/46.
XX P-PSDB: W9664.

PT New isolated LAGE-1 tumour associated nucleic acids - used to
PI develop products for the diagnosis and treatment of LAGE-1
XX associated disorders, particularly tumours

PS Claim 1: Page 53-54; 73pp; English.

XX The present sequence encodes LAGE-1 tumour associated protein (TAP).
CC The present invention also describes: (1) a method for treating a
CC subject with a disorder characterised by expression of a LAGE-1 nucleic
CC acid molecule or an expression product, comprising administering a
CC subject autologous cytolytic T cells to ameliorate the disorder, where
CC the cytolytic T cells are specific for complexes of an HLA molecule and
CC a LAGE-1 TAP or an immunogenic fragment; (2) a method for treating a
CC subject with a disorder characterised by expression of a LAGE-1 nucleic
CC acid molecule or an expression product, comprising administering a
CC LAGE-1 TAP or an immunogenic fragment to ameliorate the disorder, and
CC (3) a method for selectively enriching a population of T cells with
CC cytolytic T cells specific for a LAGE-1 TAP comprising contacting an
CC isolated population of T cells with an agent presenting a complex of a
CC LAGE-1 TAP or an immunogenic fragment and a HLA presenting molecule to
CC selectively enrich the isolated population of T cells with the cytolytic
CC T cells. The methods and products from the present invention can be used
CC for the diagnosis and treatment of LAGE-1 associated disorders.

XX Sequence 755 BP; 126 A; 234 C; 252 G; 143 I; 0 other;

Query Match 43.3%; Score 219; DB 19; Length 755;
Best Local Similarity 100.0%; Pred. No. 2.1e-50;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaatccatagatgacacgtctctgcacgaatgaagatgccctgcacacgaagaat 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 238 gaatccatagatgacacgtctctgcacgaatgaagatgccctgcacacgaagaat 297
QY 61 ggcgacacgcacgtctctgcacgaatgaagatgccctgcacacgaagaat 120
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 298 ggcgacacgcacgtctctgcacgaatgaagatgccctgcacacgaagaat 457
QY 121 gaatccatagatgacacgtctctgcacgaatgaagatgccctgcacacgaagaat 180
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 358 gaatccatagatgacacgtctctgcacgaatgaagatgccctgcacacgaagaat 417
QY 181 tctaaagacttccatgacacgtctctgcacgaatgaagatgccctgcacacgaagaat 219
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 418 tctaaagacttccatgacacgtctctgcacgaatgaagatgccctgcacacgaagaat 456

RESULT 5
ID 100150 standard; cDNA; 767 BP.
XX 100150;
XX 100150;
DI 41-JUL 2000 (first entry)
XX Human LAGE-1 splice variant, LAGE-1s encoding cDNA.
DE LAGE-1s; splice variant; CAMEL; CTL recognised Antigen on MELANOMA; CTL;
KW cytotoxic T lymphocyte; tumour associated antigen; NY-ESO-1; anticancer;
KW human; cancer; melanoma; immunotherapy; immune response; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 5'UTR 1..54
FT /*tag- a
FT CDS 54..556
FT /*tag- b

/product- "Human LAGE-1s protein"
/note "Splice variant of LAGE-1"
597..767
/*tag c

W0200023584-A1.

27-APR-2000.

15-OCT-1999; G9W0-EP07832.

16-OCT-1998; 98EP-0119584.

(BOEH) BOEHLINGER INGELHEIM INT GMBH.
(UHO-) UNIV HOSPITAL LEIDEN.

Schrier PJ, Arnould CA, Heider K, Klade C;

WPI: 2000-239685/29.

P-PSDB: Y70860.

Tumour-associated antigen useful for cancer immunotherapy is encoded by
the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA .

Example 1: Page 56-57, 73pp. English.

XX The present sequence is the cDNA encoding human LAGE-1 splice variant,
CC LAGE-1s, a tumour-specific antigen. It decodes the LAGE-1s protein from
CC open reading frame-3 (ORF) and the CAMEL protein (cytotoxic T lymphocytes
CC (CTL)-recognised Antigen on MELANOMA) from ORF-1. The LAGE-1(s/1) protein
CC from ORF-1 is identical to the CAMEL protein. LAGE-1 is not expressed in
CC healthy tissues except in testis and placenta. It also shows homology
CC with the NY-ESO-1, a melanoma specific tumour antigen. The tumour-
CC associated antigen displayed on melanoma cells is recognised by cytotoxic
CC T lymphocytes. This sequence has anticancer activity. CAMEL tumour
CC antigen and immunogenic peptides derived from it are useful for cancer
CC immunotherapy. They have the potential to induce an immune response, by
CC eliciting a CTL response. The DNA molecule is used for the construction
CC of recombinant or fusion proteins.

XX Sequence 767 BP; 138 A; 234 C; 252 G; 143 I; 0 other;

Query Match 43.3%; Score 219; DB 21; Length 767;
Best Local Similarity 100.0%; Pred. No. 2.1e-50;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaatccatagatgacacgtctctgcacgaatgaagatgccctgcacacgaagaat 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 239 gaatccatagatgacacgtctctgcacgaatgaagatgccctgcacacgaagaat 298
QY 61 ggcgacacgcacgtctctgcacgaatgaagatgccctgcacacgaagaat 120
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 299 ggcgacacgcacgtctctgcacgaatgaagatgccctgcacacgaagaat 358
QY 121 gaatccatagatgacacgtctctgcacgaatgaagatgccctgcacacgaagaat 180
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 359 gaatccatagatgacacgtctctgcacgaatgaagatgccctgcacacgaagaat 418
QY 181 tctaaagacttccatgacacgtctctgcacgaatgaagatgccctgcacacgaagaat 219
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 419 tctaaagacttccatgacacgtctctgcacgaatgaagatgccctgcacacgaagaat 457

RESULT 6
ID V50349 standard; cDNA; 217 BP.
XX V50349;
XX V50349;
XX 27 OCT-1998 (first entry)
XX Human LAGE-1 clone 1 nucleotide sequence.

RESULT: 8
 ID A61484 standard; cDNA; 752 bp.
 AC A61484;
 XX
 DT 23-OCT-2000 (first entry)
 DE cDNA encoding human oesophageal cancer-associated antigen NY-ESO-1.
 KW oesophageal cancer associated antigen; NY-ESO-1; human
 KW immunogen; oesophageal carcinoma; melanoma; ovary; testis;
 KW transmembrane domain; antibody; diagnostic marker; drug delivery target;
 KW ss.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 54..596
 FT /*tag a
 FT /product "oesophageal cancer-associated antigen,
 NY-ESO-1"
 FT misc_feature 54..600
 FT /*tag b
 FT /note "Nucleotides 54-600 are specifically claimed"
 FN OS606233-A.
 XX
 PD 30-MAY-2000.
 XX
 FE 26-JAN-1998; 980S-0013150.
 XX
 FE 03-OCT-1996; 960S-0725481.
 XX
 PA (SLOK) SLOAN KITTERLING INST CANCER RES.
 PA (CHURK) CORNELL RES FOUND INC.
 PA (LUDW) LUDWIG INST CANCER RES.
 XX
 PI Swannan M, Gure AO, Chen Y, Turaci O, Sahlin U, Pfreundschuh M;
 PI 014 L3;
 XX
 LR WP: 2000-410880/35.
 DR P PSDB: B03154.
 XX
 PI Now isolated oesophageal cancer-associated antigen useful as markers for
 PI producing antibodies and as targets for identifying abnormal
 PI conditions, e.g. infections and cancer.
 XX
 PS claim 1: Fig 4: 'pp: English.
 XX
 CC This sequence represents cDNA encoding a human oesophageal cancer-
 CC associated antigen, NY-ESO-1. This sequence was isolated from a
 CC cDNA library prepared from a specimen of well-to-moderately
 CC differentiated squamous cell cancer of the oesophagus. Expression
 CC analysis demonstrated that NY ESO-1 was expressed in oesophageal
 CC carcinoma, certain melanoma cell lines and in normal ovary and testis
 CC tissue, but not in normal colon, kidney, liver or brain tissue. Analysis
 CC of the amino acid sequence of the protein indicates that the protein has
 CC a transmembrane domain, several N-myristylation sites and
 CC O-phosphorylation sites and that it contains antigenic sequences in the
 CC N-terminal half of the protein. The antigen is useful as an immunogen
 CC when combined with an adjuvant, in both precursor and post-
 CC translationally modified forms, and may be used to generate anti-NY-ESO-1
 CC antibodies. It can also be used as a diagnostic marker for oesophageal
 CC cancer, and can be utilised as a marker for the targeted delivery of
 CC therapeutic agents to oesophageal cancer cells. It can also be used to
 CC generate diagnostic or therapeutic agents.
 XX
 SU Sequence 752 bp; 126 A; 230 C; 256 G; 140 T; 0 other;

Query Match 30.6%; Score 155; DB 21; Length 752;
 Best Local Similarity 81.7%; Pred. NO: 5,1e-33;
 Matches 179; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
 QY 1 aggtccgcatagcagctctctgagcagatgaagatgcctgcagagccacatgaag 60
 ||||| ||||| || ||||| || ||||| || ||||| ||||| ||||| ||||| |||||
 Db 239 gggccgcacagggcgcgccttcaggtctgaatgaatgctacacacacacacacac 298
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 61 gccagacagccgcctgcttcacatgacacacacacacacacacacacacacacac 120
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 246 gccac 358
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 121 ggaactggtccgcagcagctctctgagcagatgaagatgcctgcagagccacatga 180
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 259 ggaactggtccgcagcagctctctgagcagatgaagatgcctgcagagccacatga 418
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 181 tctaac 219
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 419 tctaac 457
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 RESULT 9
 D00152
 ID D00152 standard. cDNA, 752 bp.
 XX
 AC D00152;
 XX
 DT 31-JUL-2000 (first entry)
 DE Human tumour antigen, NY ESO-1 cDNA.
 XX
 KW NY ESO-1; CAMEL; CTL-recognised Antigen on Melanoma; human; cancer; (11);
 KW 2,4-dichloro-7-benzoyl-5-fluorotumour-associated antigen; (ABP-1); antitumor;
 KW melanoma; immunotherapy; immune response; ss.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT 5'UTR 1..53
 FT /*tag a
 FT CDS 54..596
 FT /*tag b
 FT /product "Human NY-ESO-1 protein"
 FT /note "Derived from open reading frame (ORF) 1"
 FT 3'UTR 597..752
 FT /*tag c
 FT 5'UTR 1..93
 FT /*tag d
 FT /note "5' UTR of ORF-2"
 FT CDS 94..270
 FT /*tag e
 FT /product "Human NY-ESO-1 short variant protein"
 FT /note "Derived from open reading frame (ORF)-2"
 FT 3'UTR 271..752
 FT /*tag f
 FT /note "3' UTR of ORF-2"
 XX
 W6205023584 A1.
 XX
 PD 27-APR-2000.
 XX
 FE 15-OCT-1999; 94W0-EP07832.
 XX
 FE 16-OCT-1998; 98EP-0119584.
 XX
 PA (PDB) BEHRINGER INGELHEIM INF GMBH.
 PA (UTHO) UNIV HOSPITAL LEIDEN.
 XX
 PI Schrier PJ, Aarnoudse CA, Heider K, Klade C;
 XX
 DR WP1: 2600-33685/29.
 DR P-PSDB: Y70862, Y70863.
 XX

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000


```
Query Match 100.0%; Score 506; DB 1; Length 1002;
Best Local Similarity 100.0%; Pred. No. 1,9e-127;
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaatcgaatgagcgtgagcgttctgcgcaggaatggaatgacccctacgaacccaaag 60
DB 250 SGGTCGGATGGCGATGAGTCTTCTGGCAGAGATGGAGGGTGGCTGGGGGCCAGAG 409
QY 61 acgaacacacacacacacacacacacacacacacacacacacacacacacacacac 120
DB 310 GAGACACACACACACACACACACACACACACACACACACACACACACACACACAC 469
QY 121 aaatcgaatgagcgtgagcgttctgcgcaggaatggaatgacccctacgaacccaaag 180
DB 370 SGGTCGGATGGCGATGAGTCTTCTGGCAGAGATGGAGGGTGGCTGGGGGCCAGAG 429
QY 181 aaatcgaatgagcgtgagcgttctgcgcaggaatggaatgacccctacgaacccaaag 240
DB 430 SGGTCGGATGGCGATGAGTCTTCTGGCAGAGATGGAGGGTGGCTGGGGGCCAGAG 489
QY 241 aaatcgaatgagcgtgagcgttctgcgcaggaatggaatgacccctacgaacccaaag 300
DB 490 SGGTCGGATGGCGATGAGTCTTCTGGCAGAGATGGAGGGTGGCTGGGGGCCAGAG 549
QY 301 aaatcgaatgagcgtgagcgttctgcgcaggaatggaatgacccctacgaacccaaag 360
DB 550 SGGTCGGATGGCGATGAGTCTTCTGGCAGAGATGGAGGGTGGCTGGGGGCCAGAG 609
QY 361 aaatcgaatgagcgtgagcgttctgcgcaggaatggaatgacccctacgaacccaaag 420
DB 610 SGGTCGGATGGCGATGAGTCTTCTGGCAGAGATGGAGGGTGGCTGGGGGCCAGAG 669
QY 421 aaatcgaatgagcgtgagcgttctgcgcaggaatggaatgacccctacgaacccaaag 480
DB 670 SGGTCGGATGGCGATGAGTCTTCTGGCAGAGATGGAGGGTGGCTGGGGGCCAGAG 729
QY 481 aaatcgaatgagcgtgagcgttctgcgcaggaatggaatgacccctacgaacccaaag 506
DB 730 SGGTCGGATGGCGATGAGTCTTCTGGCAGAGATGGAGGGTGGCTGGGGGCCAGAG 755

RESULT 2
US-08-791-495-6
Sequence 6; Application US/08791495
Patent No. 5811519
GENERAL INFORMATION:
APPLICANT: Leth, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Boon-Falleur, Thierry
TITLE OF INVENTION: LI-1 TUMOR SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows-95
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICANT'S NUMBER: US/08791495
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
```

```
US-08-791-495-1
Sequence 7; Application US/08791495
Patent No. 5811519
GENERAL INFORMATION:
APPLICANT: Leth, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Boon-Falleur, Thierry
TITLE OF INVENTION: LI-1 TUMOR SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows-95
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICANT'S NUMBER: US/08791495
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
```


117: qb_est48:*
 118: qb_est49:*
 119: qb_est50:*
 120: qb_est51:*
 121: qb_est52:*
 122: qb_est53:*
 123: qb_est54:*
 124: qb_est55:*
 125: qb_est56:*
 126: qb_est57:*
 127: qb_est58:*
 128: qb_est59:*
 129: qb_est60:*
 130: qb_est61:*
 131: qb_est62:*
 132: qb_est63:*
 133: qb_est64:*
 134: qb_est65:*
 135: qb_est66:*
 136: qb_est67:*
 137: qb_est68:*
 138: qb_est69:*
 139: qb_est70:*
 140: qb_est71:*
 141: qb_est72:*
 142: qb_est73:*
 143: qb_est74:*
 144: qb_est75:*
 145: qb_est76:*
 146: qb_est77:*
 147: qb_est78:*
 148: qb_est79:*
 149: qb_est80:*
 150: qb_est81:*
 151: qb_est82:*
 152: qb_est83:*
 153: qb_est84:*
 154: qb_est85:*
 155: qb_est86:*
 156: qb_est87:*
 157: qb_est88:*
 158: qb_est89:*
 159: qb_est90:*
 160: qb_est91:*
 161: qb_est92:*
 162: qb_est93:*
 163: qb_est94:*
 164: qb_est95:*
 165: qb_est96:*
 166: qb_est97:*
 167: qb_est98:*
 168: qb_est99:*
 169: qb_est100:*
 170: qb_est101:*
 171: qb_est102:*
 172: qb_est103:*
 173: qb_est104:*
 174: qb_est105:*
 175: qb_est106:*
 176: qb_est107:*
 177: qb_est108:*
 178: qb_est109:*
 179: qb_est110:*
 180: qb_est111:*
 181: qb_est112:*
 182: qb_est113:*
 183: qb_est114:*
 184: qb_est115:*
 185: qb_est116:*
 186: qb_est117:*
 187: qb_est118:*
 188: qb_est119:*
 189: qb_est120:*

190: em_qss_pln1:*
 191: em_qss_pln2:*
 192: em_qss_pro:*
 193: em_qss_rod1:*
 194: em_qss_rod2:*
 195: em_qss_rod3:*
 196: em_qss_rod4:*
 197: em_qss_rod5:*
 198: em_qss_vrt1:*
 199: em_qss_vrt2:*
 200: em_qss_vrt3:*
 201: qb_qss1:*
 202: qb_qss2:*
 203: qb_qss3:*
 204: qb_qss4:*
 205: qb_qss5:*
 206: qb_qss6:*
 207: qb_qss7:*
 208: qb_qss8:*
 209: qb_qss9:*
 210: qb_qss10:*
 211: qb_qss11:*
 212: qb_qss12:*
 213: qb_qss13:*
 214: qb_qss14:*
 215: qb_qss15:*
 216: qb_qss16:*
 217: qb_qss17:*
 218: qb_qss18:*
 219: qb_qss19:*
 220: qb_qss20:*
 221: qb_qss21:*
 222: qb_qss22:*
 223: qb_qss23:*
 224: qb_qss24:*
 225: qb_qss25:*
 226: qb_qss26:*
 227: qb_qss27:*
 228: qb_qss28:*
 229: qb_qss29:*
 230: qb_qss30:*
 231: qb_qss31:*
 232: qb_qss32:*
 233: qb_qss33:*
 234: qb_qss34:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|-----|----------|---------------------|
| 1 | 416.6 | 82.3 | 699 | 165 | BE274123 | BE274123 601120646 |
| 2 | 345.2 | 58.2 | 596 | 166 | BE390259 | BE390259 601285131 |
| 3 | 327.9 | 61.8 | 559 | 155 | BE276194 | BE276194 601144462 |
| 4 | 313.2 | 61.9 | 739 | 166 | BE391213 | BE391213 601285820 |
| 5 | 279.4 | 55.2 | 809 | 166 | BE408892 | BE408892 601337769 |
| 6 | 225.2 | 44.5 | 674 | 166 | BE388562 | BE388562 601281855 |
| 7 | 203.2 | 40.2 | 645 | 166 | BE410697 | BE410697 601360763 |
| c 8 | 185.2 | 36.6 | 852 | 166 | BE385880 | BE385880 601275948 |
| 9 | 175.4 | 31.7 | 511 | 166 | BE387918 | BE387918 601282166 |
| 10 | 156.8 | 31.0 | 483 | 166 | BE410952 | BE410952 601303669 |
| c 11 | 59.4 | 11.7 | 359 | 17 | AL218223 | AL218223 qb2493.x |
| 12 | 56.4 | 11.1 | 534 | 31 | AV663843 | AV663843 AV663843 |
| 13 | 54 | 10.7 | 639 | 166 | BE387002 | BE387002 601274781 |
| c 14 | 51.6 | 10.2 | 925 | 229 | CNS0091P | AL053013 Dr osoph11 |
| 15 | 50.4 | 10.0 | 925 | 229 | CNS0091P | AL053013 Dr osoph11 |
| c 16 | 49.6 | 9.8 | 498 | 111 | AW161857 | AW161857 au88d01.x |
| 17 | 49.6 | 9.8 | 559 | 111 | AW157381 | AW157381 au88d01.x |
| c 18 | 49.6 | 9.8 | 714 | 31 | AV703867 | AV703867 AV703867 |

GenCore version 4.5
Copyright (c) 1993-2000 CompuGen Ltd.

0M nucleotide nucleotide search, using sw mode 1

Run date: May 1, 2001, 09:51:24 ; Search time: 2512.89 Seconds
(without alignments)
123,247 Million cell updates/sec

File: us_09_341_829a_4_copy_488_508
Perfect score: 21
Sequence: 1-annuunp-qqetuuqquuqaa_21

Scoring table: IDENTITY, 80%
Gapop 10.0 ; Gapext 1.0

Searches: 12082.0 seqs, 737929662 residues
Total number of hits satisfying chosen parameters: 2566470

Minimum hit seq length: 0

Maximum hit seq length: 200000000

Post processing: Maximum Match 100%
Listed first 4% summaries

Database:

1: db_bact1*
2: db_bact2*
3: db_bact3*
4: db_bact4*
5: db_bact5*
6: db_bact6*
7: db_bact7*
8: db_bact8*
9: db_bact9*
10: db_bact10*
11: db_bact11*
12: db_bact12*
13: db_bact13*
14: db_bact14*
15: db_bact15*
16: db_bact16*
17: db_bact17*
18: db_bact18*
19: db_bact19*
20: db_bact20*
21: db_bact21*
22: db_bact22*
23: db_bact23*
24: db_bact24*
25: db_bact25*
26: db_bact26*
27: db_bact27*
28: db_bact28*
29: db_bact29*
30: db_bact30*
31: db_bact31*
32: db_bact32*
33: db_bact33*
34: db_bact34*
35: db_bact35*
36: db_bact36*
37: db_bact37*
38: db_bact38*
39: db_bact39*
40: db_bact40*
41: db_bact41*
42: db_bact42*
43: db_bact43*

44: db_bact44*
45: db_bact45*
46: db_bact46*
47: db_bact47*
48: db_bact48*
49: db_bact49*
50: db_bact50*
51: db_bact51*
52: db_bact52*
53: db_bact53*
54: db_bact54*
55: db_bact55*
56: db_bact56*
57: db_bact57*
58: db_bact58*
59: db_bact59*
60: db_bact60*
61: db_bact61*
62: db_bact62*
63: db_bact63*
64: db_bact64*
65: db_bact65*
66: db_bact66*
67: db_bact67*
68: db_bact68*
69: db_bact69*
70: db_bact70*
71: db_bact71*
72: db_bact72*
73: db_bact73*
74: db_bact74*
75: db_bact75*
76: db_bact76*
77: db_bact77*
78: db_bact78*
79: db_bact79*
80: db_bact80*
81: db_bact81*
82: db_bact82*
83: db_bact83*
84: db_bact84*
85: db_bact85*
86: db_bact86*
87: db_bact87*
88: db_bact88*
89: db_bact89*
90: db_bact90*
91: db_bact91*
92: db_bact92*
93: db_bact93*
94: db_bact94*
95: db_bact95*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | % Match | | Description |
|------------|-------|-------------|---------|-----|--------------------|
| | | | Length | Hit | |
| 1 | 21 | 100.0 | 994 | 9 | AX024689 Sequence |
| 2 | 21 | 100.0 | 994 | 92 | AF012845 Homo sapi |
| 3 | 21 | 100.0 | 1002 | 9 | AK042555 Sequence |
| 4 | 21 | 100.0 | 1002 | 92 | AJ224040 Homo sapi |
| 5 | 21 | 100.0 | 1002 | 91 | AJ224094 Homo sapi |
| 6 | 21 | 100.0 | 112874 | 78 | AF277415 Homo sapi |
| 7 | 17.8 | 84.8 | 2931 | 93 | U10694 Human MAGE- |
| 8 | 17.8 | 84.8 | 40438 | 88 | AC087410 Homo sapi |
| 9 | 17.8 | 84.8 | 43927 | 93 | U65668 Human Xq28 |
| 10 | 17.8 | 84.8 | 73360 | 93 | U66083 Human cont1 |
| 11 | 17.8 | 84.8 | 111560 | 87 | AC016940 Homo sapi |

| | | | | | | |
|----|------|------|--------|----|-----------|--------------------|
| 12 | 17.8 | 84.8 | 118440 | 87 | AC016939 | Homo sapi |
| 13 | 17.8 | 84.8 | 214894 | 70 | AC025751 | Mus muscu |
| 14 | 17.4 | 82.9 | 4900 | 14 | PC11PARG | X54257 P. chrysosp |
| 15 | 17.4 | 82.9 | 10522 | 1 | AE034514 | AF004534 Pseudomon |
| 16 | 17.4 | 82.9 | 339681 | 2 | AF004009 | AF004009 Mesorhizo |
| 17 | 17.4 | 82.9 | 345783 | 2 | AF004001 | AF004001 Mesorhizo |
| 18 | 16.8 | 80.0 | 58280 | 3 | MTV014 | AL021646 Myoblaste |
| 19 | 16.8 | 80.0 | 118947 | 60 | AC006321 | AC006321 Homo sapi |
| 20 | 16.8 | 80.0 | 124040 | 83 | CNS010SLJ | AL121808 Homo sapi |
| 21 | 16.8 | 80.0 | 124337 | 92 | HSRM16C9 | AL121825 Human DNA |
| 22 | 16.8 | 80.0 | 132698 | 60 | AC005051 | AC005051 Homo sapi |
| 23 | 16.8 | 80.0 | 167457 | 72 | AC037478 | AC037478 Homo sapi |
| 24 | 16.8 | 80.0 | 190744 | 85 | AC005038 | AC005038 Homo sapi |
| 25 | 16.8 | 80.0 | 190846 | 86 | AC006372 | AC006372 Homo sapi |
| 26 | 16.8 | 80.0 | 193038 | 61 | AC008742 | AC008742 Homo sapi |
| 27 | 16.8 | 80.0 | 213355 | 63 | AC012451 | AC012451 Homo sapi |
| 28 | 16.4 | 78.1 | 964 | 94 | MEAF27R | U07411 Mouse mRNA |
| 29 | 16.4 | 78.1 | 964 | 94 | MEAF27R | U07411 Mouse mRNA |
| 30 | 16.4 | 78.1 | 1464 | 94 | AF093403 | AF093403 Mus muscu |
| 31 | 16.4 | 78.1 | 2436 | 94 | AF288694 | AF288694 Mus muscu |
| 32 | 16.4 | 78.1 | 3528 | 88 | AF014643 | AF014643 Homo sapi |
| 33 | 16.4 | 78.1 | 4538 | 88 | AF147709 | AF147709 Homo sapi |
| 34 | 16.4 | 78.1 | 23610 | 62 | AC011805 | AC011805 Homo sapi |
| 35 | 16.4 | 78.1 | 41604 | 93 | HSAY18000 | U011932 Homo sapi |
| 36 | 16.4 | 78.1 | 126138 | 92 | HSAY18000 | U011932 Homo sapi |
| 37 | 16.4 | 78.1 | 134036 | 86 | AC007845 | AC007845 Homo sapi |
| 38 | 16.4 | 78.1 | 136030 | 90 | AL359815 | AL359815 Homo sapi |
| 39 | 16.4 | 78.1 | 152137 | 90 | AL357252 | AL357252 Homo sapi |
| 40 | 16.4 | 78.1 | 157023 | 92 | AL513129 | AL513129 Homo sapi |
| 41 | 16.4 | 78.1 | 171878 | 85 | AC004882 | AC004882 Homo sapi |
| 42 | 16.4 | 78.1 | 180187 | 68 | AC024093 | AC024093 Homo sapi |
| 43 | 16.4 | 78.1 | 183920 | 80 | AL459510 | AL459510 Homo sapi |
| 44 | 16.4 | 78.1 | 194982 | 81 | AL451069 | AL451069 Homo sapi |
| 45 | 16.4 | 78.1 | 201373 | 75 | AC074767 | AC074767 Mus muscu |

| | |
|------------|---|
| RESULT | 1 |
| LOCUS | AX024689 |
| DEFINITION | Sequence 5 from Patent EP1001022. |
| ACCESSION | AX024689 |
| VERSION | AX024689.1 GI:101941775 |
| KEYWORDS | |
| SOURCE | human. |
| ORGANISM | Homo sapiens |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eupleurostomi |
| AUTHORS | 1 (bases 1 to 994) |
| TITLE | Aarnoudse, C.A. and Schrier, P. |
| JOURNAL | Cell, an alternative translation product of the tumour antigen |
| FEATURES | Location/Qualifiers |
| Source | 1..994 |
| gene | 1..994 |
| CDS | 56..688 |
| 5'UTR | 56..688 |
| CDS | 56..688 |

| | | | | |
|------------|-------|-------|-------|-------|
| BASE COUNT | 184 a | 295 c | 334 g | 180 t |
| ORIGIN | | | | |

| | | | | |
|-----------------------|----------------|---------------------------|----------|------------|
| Query Match | 100.0% | Score 21 | DB 9 | Length 994 |
| Best Local Similarity | 100.0% | Prod. No. 26 | | |
| Matches 21 | Conservative 9 | Mismatches 0 | Indels 0 | Gaps 0 |
| QY | 1 | aaagaaagcactatgacacata 21 | | |
| DB | 479 | AGGGAAGGATATGGGCGGATG 499 | | |

| | |
|------------|---|
| RESULT | 2 |
| LOCUS | HSAL012835 |
| DEFINITION | Homo sapiens mRNA for LAGE-1L and CTL-recognized antigen on melanoma (CAMEL). |
| ACCESSION | U012835 |
| KEYWORDS | CAMEL gene; CTL recognized antigen on melanoma (CAMEL); LAGE 1L gene. |
| SOURCE | human. |
| ORGANISM | Homo sapiens |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. |
| AUTHORS | 1 (bases 1 to 994) |
| TITLE | Aarnoudse, C.A. and Schrier, P. |
| JOURNAL | Cell, an alternative translation product of the tumour antigen |
| FEATURES | Location/Qualifiers |
| Source | 1..994 |
| gene | 1..994 |
| CDS | 56..688 |
| 5'UTR | 56..688 |
| CDS | 56..688 |

| | | | | |
|------------|-------|-------|-------|-------|
| BASE COUNT | 184 a | 295 c | 434 g | 180 t |
| ORIGIN | | | | |

Garreil, J. H., Correll, L. L., Guevara, W., Harris, K., Brannan, J., Hodgson, A., Hoque, M., Holloway, C., Hosak, H., Jackson, J. B., Jackson, L., Jia, Y., Jones, M., Kelly, S., Kondejowski, N., Kud, Y., Kwart, C., Leal, B., Li, Z., Lichtarge, O., Liu, J., Liu, W., Logan, O., Lovado, R. J., Lu, J., Lurie, R., Martin, R., Martinez, C., McLeod, M. P., Mei, G., Moran, M., Morris, S., Rast, J., Reichen, A., Ripstein, R., Nuyten, N., Nguyen, S., Ouellet, G., Pamrell, J. K., Parish, A., Paxton, S., Payne, B., Perez, L., Pichler, L., Quiles, M., Rashid, N. D., Reiter, D., Rives, M., Samuel, S., Say, J., Scherer, S., Shah, E., Shen, H., Simon, W., Sparks, A., Stamps, A., Suvarant, R., Taper, P., Taylor, T., Vasquez, I., Vinson, P., Vo, Q., Wahbeh, M., Watlington, S., Weinstein, G., Weinstein, I. P., Willington, A., Worley, E., Wren, J., Wrenston, G., Yu, W., Zhou, X., Nelson, D., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 111560)
Worley, K.C.
Direct Submission
Submitted (09-DEC-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 111560)
Worley, K.C.
Direct Submission
Submitted (01-APR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
on Apr 1, 2000 this sequence version replaced gi:7007873.
INFO: genome browser: <http://www.hgsc.bcm.tmc.edu>
help: <http://www.hgsc.bcm.tmc.edu>

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:

STES are identified using either CHASER (v. 7.54), GCG, or a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for human and mouse sequences.

Genes and region of sequence similarity are identified by BLAST (Proc. Acad. Sci. 93:389-392) similarity (expect < 1e-35) to the EST and cDNA sequences. Genes are annotated at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 4 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at http://www.ncbi.nlm.nih.gov/sra/seq/seq_quality.html.

QUALSTAT-REPORT-----

----- Summary Statistics -----
Contig length: 111560
Phrap values in estimate: 111560
Average error rate (RCM Phrap estimate): 0.000108506
Fraction of Phrap values less than 40: 0.0136415

Number of consensus changed edits:
Number of N's in consensus:

2
0

----- Consensus changing edits -----
Feature of Global Context Edited Context
35684 aaaaaaaaaa(n)acacaaataa aaaaaaaaaa(a)aaacaaataa
82852 aaaaaaaaaa(n)aaaaaaataa aaaaaaaaaa(a)aaaaaaataa

----- Distribution of Quality < 40 Bases -----

| | 1000 | 900 | 800 | 700 | 600 | 500 | 400 | 300 | 200 | 100 | 0 | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 |
|-------------------|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|---|---|----|----|----|----|----|----|----|
| # bases | 1000 | 900 | 800 | 700 | 600 | 500 | 400 | 300 | 200 | 100 | 0 | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 |
| Phrap Value Range | | | | | | | | | | | | | | | | | | | |

Version: 1.01 qxf0.

FEATURES

Location/Qualifiers
1..111560

Accession: "Homo sapiens"

Chromatid: "taxon:9606"

Chromosome: "Xq28"

Clone: "RP5-937E21"

Complement(883..1013)

/rpt_family "LIMB3"

1020..1094

/rpt_family "L2"

Complement(1765..2081)

/rpt_family "L2"

Complement(2420..3114)

/rpt_family "LIM-2"

Complement(3115..3248)

/rpt_family "LIM-3"

3252..3613

/rpt_family "LIMCB"

3619..4148

/rpt_family "LIM-5"

4931..5391

/rpt_family "L1"

Complement(7992..8022)

/rpt_family "MER50"

8033..8055

/rpt_family "(CAA)n"

Complement(8056..8190)

/rpt_family "MER50"

Complement(8191..8547)

/rpt_family "HEP18"

Complement(9548..8530)

/rpt_family "MER50"

Complement(9549..974)

/rpt_family "MER50"

Complement(8784..8835)

/rpt_family "MER50"

8892..90056

/note "region similar to HUMMAG1A Human antigen (MAGE-1)

gene: M77481"

/note "region: HUMMAG1A Human antigen (MAGE-1) gene"

9694..9887

/note "region similar to Hs4552485 Human MAGE-10 antigen

(MAGE10) gene: U10685"


```

cDS
/contig="PA0794"
/complement(6582..9188)
/contig="PA0794"
/codon_start=1
/transl_table=11
/product="probable ascorbate hydratase"
/protein_id "AAG04184.1"
/db_xref="GI:5946685"
/translation "MSSARRKPTLESTFLNEGAREAVEAIGCTAVAKLEYTSKVLADN
LVKQCTPATLEASTPOLVERPDLDFWYFAPVVCILIGCTAVVLAAGLKLALAKG
GDAQNVVAVVULIVDHSIAVECCGYDEFAAKNAIEDRNEHRPHFLDMTKQAFK
NVIVIPGNCIMQINLEKNSVIOQRKQVAFEDTCVGDSDHPHVDALGVIAUGGG
LEAHVHELKASRHLDFIVGPEISGRPGGTHIIVIALIEFLKQKVGAVLEFY
GESASLLIDRATILNMAPEYGATAAMPAIDQQTIDVLRITGRDEQVALVEAYRT
AGLWASLVIAEYERVLKEFDLSNVNMACPSNPHAKVAISELAAGTLAGNLEKARAE
FAKLMFGAVIIAATISCTINSNPNVIAAGLLARNADRLGVKPKWKTSLAQSK
VVTHYLEALILHLEALIGGVVAVACTSCNGMSGALDPAIQREIVKRLYATAVLG
NRNDEGLHYAKQAFLEPIWAVAYIACTTFEDIEVIGVIGREIRLEDLPSD
RELDVVRRAVKEQERQVITDFDITGGEREKVIDIYAWRPTSTYIRRTPTWEGALA
GERTLRMRVLAVLPNITDHLSPNAILADSAAGEXIAKMGLEPEEFNSYATIRGID
HLLTAQRATFAMPKLFNEVNRNAGSVKQGLARVEFGKVMRMMAIEITYMERKQPLI
IVACADYGGSSPDAAKGVRIAGVEAVAGGFERIIRTNLIGMGLVLPLEKPKITRL
TIGLDGSETHDVLGARRPRADITLVIRHRDGERLEVPVTCRLDSIEEVSITYEAGGVIG
READPLEAGA"
/complement(9319..10446)
/contig="prpc"
/contig="PA0795"
/complement(9319..10446)
/contig="prpc"
/codon_start=1
/transl_table=11
/product="citrate synthase 2"
/protein_id "AAG04184.1"
/db_xref="GI:5946686"
/translation "MAEAKVLSGAGLRQVATATSTVQEGGALTYGYGVVRLAA
AAIFERVAVILAYGELFNKQGLAVLAKLQQRGEPQALKEVLEETADAHMVMHRI
GASVLTLEHLEISPDQORVADRLAAPPATMYWYRTREGQRIDCNSDEPTIGGHP
LALLRQKKISLHVAVNVSILLYAHIEHPNASTPTARWCATSLDLYSCVIGATGSLR
GPLHGAREAAAMELIERFSGEATAELIKMIFETKIMGEFHAIFYLSDFPNEVIRG
WPKQADEGVKRVLEAVSAGADKIMWEKRLFPNALDFHASAYHFMGLPTKLEPTTFV
CSRTSGWIAHVPFQRANNRIHPSAEYTGVEQRAFPVPLEQR"
BASE COUNT 1529 a 3412 c 3723 g 1858 t
ORIGIN

```

```

Query Match: 82.9%; Score 17.4; DB 1; Length 10522;
Best Local Similarity 94.7%; Pred. No. 6.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 3 qraaardqctqacqatq 21
  ||| ||| ||| ||| ||| |||
DB 6544 GGAAGGCGCTGGGGGGCTG 3626

```

Search completed: May 1, 2001, 03:51:35
Job time: 11628 sec

PA (B-EH) BOEHRINGER INGELHEIM INT GMBH.
 XX (UYHE-) UNIV HOSPITAL LEIDEN.
 PI Schrier PL, Aarnoudse CA, Heider K, Klade C;
 XX
 DR WPI: 2000-039685/29.
 DR P-PSDB: Y70861.
 XX
 XX Tumor associated antigen useful for cancer immunotherapy is encoded by
 PI the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA -
 XX
 PS Example 1: Page 58-59; 73pp; English.
 XX
 CC The present sequence is the cDNA encoding human LAGE-1 unspliced variant,
 CC LAGE-1L, a tumor-specific antigen. It decodes the LAGE-1L protein from
 CC open reading frame-1 (ORF) and the CAMPE protein (cytotoxic T lymphocyte
 CC CTL) recognized antigen (cancer-associated protein). The LAGE-1L protein
 CC from ORF 1 is identical to the CAMPE protein. LAGE-1 is not expressed in
 CC healthy tissues except in testis and placenta. It also shows homology
 CC with the NY-ESO-1, a melanoma specific tumor antigen. The tumor-
 CC associated antigen displayed on melanoma cells is recognized by cytotoxic
 CC T lymphocytes. This sequence has anticancer activity. CAMPE tumor
 CC antigen and immunogenic peptides derived from it are useful for cancer
 CC immunotherapy. They have the potential to induce an immune response, by
 CC eliciting a CTL response. The DNA molecule is used for the construction
 CC of recombinant or fusion proteins.
 XX
 SQ Sequence 993 BP; 184 A; 295 C; 444 G; 180 T; 0 other;

Query Match 100.0%; Score 21; DB 21; Length 993;
 Best Local Similarity 100.0%; Pred. No. 1,2;
 Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaaaaagcgcctqqcqaatg 21
 bb 479 aaaaaagcgcctqqcqaatg 499

RESULT 2
 V0346
 ID V50346 standard: cDNA; 1002 BP.
 XX
 AC V50346;
 XX
 DI 27-OCT-1998 (first entry)
 XX
 DE Human LAGE-1 clone 2 nucleotide sequence.
 XX
 KW Human LAGE-1: tumour associated protein; LL 1, diagnosis; tumour ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH CDS 65..647
 FH /start a
 FH /product "LAGE-1"
 XX
 FN W09832855.A1.
 XX
 PD 30-JUL-1998.
 XX
 PF 27 JAN-1998; 98WO-0501445.
 XX
 PE 27-JAN-1997; 97OS-0791495.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Beau-Falloux T, De Smet C, Godetaine D, Letho B;
 PI Lucas S;
 XX
 DR WPI: 1998-427951/30.
 DR P-PSDB: W69663.

XX New isolated LAGE-1 tumour associated nucleic acids used to
 PI develop products for the diagnosis and treatment of LAGE-1
 PI associated disorders, particularly tumours
 XX
 PS Claim 1: Page 50-52; 73pp; English.
 XX
 CC The present sequence encodes LAGE-1 tumour associated protein (LAP).
 CC The present invention also describes: (1) a method for treating a
 CC subject with a disorder characterised by expression of a LAGE-1 nucleic
 CC acid molecule or an expression product, comprising administering a
 CC subject autologous cytolytic T cells to ameliorate the disorder, where
 CC the cytolytic T cells are specific for complexes of an HLA molecule and
 CC a LAGE-1 TAP or an immunogenic fragment; (2) a method for treating a
 CC subject with a disorder characterised by expression of a LAGE-1 nucleic
 CC acid molecule or an expression product, comprising administering a
 CC LAGE-1 TAP or an immunogenic fragment to ameliorate the disorder; and
 CC (3) a method for selectively enriching a population of T cells with
 CC cytolytic T cells specific for a LAGE-1 TAP comprising contacting an
 CC isolated population of T cells with an agent presenting a complex of a
 CC LAGE-1 TAP or an immunogenic fragment and a HLA presenting molecule to
 CC selectively enrich the isolated population of T cells with the cytolytic
 CC T cells. The methods and products from the present invention can be used
 CC for the diagnosis and treatment of LAGE-1 associated disorders.
 XX
 SQ Sequence 1002 BP; 179 A; 302 C; 436 G; 185 T; 0 other;

Query Match 100.0%; Score 21; DB 19; Length 1002;
 Best Local Similarity 100.0%; Pred. No. 1,2;
 Matches 21: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaaaaagcgcctqqcqaatg 21
 bb 488 aaaaaagcgcctqqcqaatg 508

RESULT 3
 C68127
 ID C68127 standard: cDNA; 1769 BP.
 XX
 AC C68127;
 XX
 DI 20 FEB-2001 (first entry)
 XX
 DE Human secreted protein cDNA sequence #47.
 XX
 KW Cytostatic; immunosuppressive; neutropenic; neutroprotective; antiviral;
 KW antifibrotic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW oncolytic; antihypertensive; antibacterial; antifungal; antiparasitic;
 KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human, secreted protein; SS.
 XX
 OS Homo sapiens.
 XX
 FN W1200058335.A1.
 XX
 PD 05-OCT-2000.
 XX
 PF 22 MAR-2000; 2000WO-0507534.
 XX
 PE 26-MAR-1999; 99US-0126598.
 XX
 PD 22 FEB-1999; 99US-0171504.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSK/) ROSEN C A.
 XX
 PI Rosen CA, Ruben SM, Komatsu S G;
 DR WPI: 2000-611702/58.
 DR P-PSDB: B37394.
 XX

14. Nucleic acids encoding human secreted proteins, used to treat, prevent, ameliorate or diagnose conditions such as cancer, and autoimmune diseases e.g. arthritis

XX
L2S
Columb; Papyrus 447-448; 487top; Perilous

The investigation relates to the isolation of genes (58891-59127) encoding 47 human secreted proteins (57448-57764); the genes can be used to generate fusion proteins by linking to the gene for the human immunoglobulin G κ chain (576072) for increasing the stability of the fusion protein as compared to the human protein only. The genes and proteins are useful for preventing, amelioration or treating medical conditions, e.g. by protein or gene therapy; the genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, and antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer, and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological disorders e.g. cerebral anaemia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Summary: 176,944; 6,11 A; 47,9 C; 417 G; 46,2 F; 0 other;

Country Match. 77.18; Score 16.2; LBS 21; Length 1769;
Best Total Similarity 86.7%; Prod. No. 1.4002;

[illegible]

REF: 4
 2005/5/17
 [1] Z. 1983 3451 and 1983 3452
 CNA: 480 180

XXV. 1111

XX

inflammatory, cardiovascular, disease, and immune, and inflammatory;
inflammatory, heart, liver, neuroprotective, cardiovascular, hepatocellular;
and cardiovascular, in the brain, and protective, neurological disorder,
neurological disorder, development of disorder, after loss of tissue;
cardiovascular disease, acquired immune deficiency syndrome, anemia;
stroke, heart disease, infection, Alzheimer's disease, schizophrenia;
parkinson's disease, bipolar disorder, evolutionary defect,
immune, and other, 50.

100

$$W_1 = \frac{1}{2} \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \right)^2 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2} I_2$$

XX

$$X \times X = [X] \cup [X] + [X]$$

| α | β | γ | δ | ϵ | ζ | η | θ | ι | κ | λ | μ | ν | ξ | \omicron | π | ρ | σ | τ | υ | ϕ | χ | ψ | ω | |
|----------|---------|----------|----------|------------|---------|--------|----------|---------|----------|-----------|-------|-------|-------|------------|-------|--------|----------|--------|------------|--------|--------|--------|----------|----|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |

{ } III. I'gag; 'yyl'; '()' g'g'gh'.

1964, 1965, 1966, 1967, 1968, 1969, 1970, 1971, 1972, 1973, 1974, 1975, 1976, 1977, 1978, 1979, 1980, 1981, 1982, 1983, 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 26

VIA (N.Y.) N.Y.E. PHARM INC.

PI Lai P., Tang YL, Chengong GA, Guo YR, Guo YR, Guo YR, Guo YR;
 PI Akerblom E, Au-Yang J, Yoo H, Patterson C, Keshy R, Hillman J;
 PI Bandman O;

XX
DR
DR

XX New human shared peptides containing proteins useful in treatment, prevention and diagnosis of e.g. cancer, inflammation and cardiovascular disease.

XX
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536

258139 1G7624; encode Y87247 to Y87247 which represent the human signal peptide containing proteins HSP61 to HSP104. HSPs have anti-infective, antimicrobial, anti-inflammatory, neutrophil, hepatocyte, neuroprotective, cardiovascular and osteolytic activities, and can be used in gene therapy. HSPs can be used to treat or prevent disorders associated with decreased activity or function of HSP. Antagonists of HSP are used to treat or prevent disorders associated with increased activity or function of HSP. Such disorders include cell proliferation (including cancer), inflammation, cardiovascular, neurological, reproductive or developmental disorders, (e.g., arteriosclerosis, cirrhosis, psoriasis), acquired immune deficiency syndrome, anemia, asthma, Crohn's disease, microbial or other infections, congestive or ischemic heart disease, Alzheimer's, Parkinson's or Huntington's diseases, schizophrenia, regulatory defects, muscular dystrophy). HSP nucleic acids can be used for the recombinant production of HSP. For detecting HSP in standard hybridisation and amplification assays (for diagnosis and monitoring) in gene therapy, as antisense, triplex-forming or ribozyme therapeutics, for detecting related sequences of genetic variations, and for chromosomal mapping, HSP are also used to raise specific antibodies (Ab) and to screen for agonists and antagonists (potential therapeutic agents). Ab are used to diagnose, or monitor, HSP-related diseases (in usual immunoassays), as therapeutic antagonists, in competitive drug screens, and for purification of HSP from natural sources.

Sequence 480 bp: 113 A; 128 C; 128 G; 101 T; 1 other.

| | | | | | | | |
|-----------------------|--------|-----------|---------|----|----|--------|-----|
| Query Match | 76.28 | Score | 16 | DB | 21 | Length | 480 |
| Best Local Similarity | 100.08 | Prod. No. | 1.7e+02 | | | | |

16

165 AAGGAGGGCTGGGC 150

RESULTS

11/195X
12/195X
13/195X
14/195X
15/195X
16/195X
17/195X
18/195X
19/195X
20/195X
21/195X
22/195X
23/195X
24/195X
25/195X
26/195X
27/195X
28/195X
29/195X
30/195X
31/195X
32/195X
33/195X
34/195X
35/195X
36/195X
37/195X
38/195X
39/195X
40/195X
41/195X
42/195X
43/195X
44/195X
45/195X
46/195X
47/195X
48/195X
49/195X
50/195X
51/195X
52/195X
53/195X
54/195X
55/195X
56/195X
57/195X
58/195X
59/195X
60/195X
61/195X
62/195X
63/195X
64/195X
65/195X
66/195X
67/195X
68/195X
69/195X
70/195X
71/195X
72/195X
73/195X
74/195X
75/195X
76/195X
77/195X
78/195X
79/195X
80/195X
81/195X
82/195X
83/195X
84/195X
85/195X
86/195X
87/195X
88/195X
89/195X
90/195X
91/195X
92/195X
93/195X
94/195X
95/195X
96/195X
97/195X
98/195X
99/195X
100/195X
101/195X
102/195X
103/195X
104/195X
105/195X
106/195X
107/195X
108/195X
109/195X
110/195X
111/195X
112/195X
113/195X
114/195X
115/195X
116/195X
117/195X
118/195X
119/195X
120/195X
121/195X
122/195X
123/195X
124/195X
125/195X
126/195X
127/195X
128/195X
129/195X
130/195X
131/195X
132/195X
133/195X
134/195X
135/195X
136/195X
137/195X
138/195X
139/195X
140/195X
141/195X
142/195X
143/195X
144/195X
145/195X
146/195X
147/195X
148/195X
149/195X
150/195X
151/195X
152/195X
153/195X
154/195X
155/195X
156/195X
157/195X
158/195X
159/195X
160/195X
161/195X
162/195X
163/195X
164/195X
165/195X
166/195X
167/195X
168/195X
169/195X
170/195X
171/195X
172/195X
173/195X
174/195X
175/195X
176/195X
177/195X
178/195X
179/195X
180/195X
181/195X
182/195X
183/195X
184/195X
185/195X
186/195X
187/195X
188/195X
189/195X
190/195X
191/195X
192/195X
193/195X
194/195X
195/195X
196/195X
197/195X
198/195X
199/195X
200/195X
201/195X
202/195X
203/195X
204/195X
205/195X
206/195X
207/195X
208/195X
209/195X
210/195X
211/195X
212/195X
213/195X
214/195X
215/195X
216/195X
217/195X
218/195X
219/195X
220/195X
221/195X
222/195X
223/195X
224/195X
225/195X
226/195X
227/195X
228/195X
229/195X
230/195X
231/195X
232/195X
233/195X
234/195X
235/195X
236/195X
237/195X
238/195X
239/195X
240/195X
241/195X
242/195X
243/195X
244/195X
245/195X
246/195X
247/195X
248/195X
249/195X
250/195X
251/195X
252/195X
253/195X
254/195X
255/195X
256/195X
257/195X
258/195X
259/195X
260/195X
261/195X
262/195X
263/195X
264/195X
265/195X
266/195X
267/195X
268/195X
269/195X
270/195X
271/195X
272/195X
273/195X
274/195X
275/195X
276/195X
277/195X
278/195X
279/195X
280/195X
281/195X
282/195X
283/195X
284/195X
285/195X
286/195X
287/195X
288/195X
289/195X
290/195X
291/195X
292/195X
293/195X
294/195X
295/195X
296/195X
297/195X
298/195X
299/195X
300/195X
301/195X
302/195X
303/195X
304/195X
305/195X
306/195X
307/195X
308/195X
309/195X
310/195X
311/195X
312/195X
313/195X
314/195X
315/195X
316/195X
317/195X
318/195X
319/195X
320/195X
321/195X
322/195X
323/195X
324/195X
325/195X
326/195X
327/195X
328/195X
329/195X
330/195X
331/195X
332/195X
333/195X
334/195X
335/195X
336/195X
337/195X
338/195X
339/195X
340/195X
341/195X
342/195X
343/195X
344/195X
345/195X
346/195X
347/195X
348/195X
349/195X
350/195X
351/195X
352/195X
353/195X
354/195X
355/195X
356/195X
357/195X
358/195X
359/195X
360/195X
361/195X
362/195X
363/195X
364/195X
365/195X
366/195X
367/195X
368/195X
369/195X
370/195X
371/195X
372/195X
373/195X
374/195X
375/195X
376/195X
377/195X
378/195X
379/195X
380/195X
381/195X
382/195X
383/195X
384/195X
385/195X
386/195X
387/195X
388/195X
389/195X
390/195X
391/195X
392/195X
393/195X
394/195X
395/195X
396/195X
397/195X
398/195X
399/195X
400/195X
401/195X
402/195X
403/195X
404/195X
405/195X
406/195X
407/195X
408/195X
409/195X
410/195X
411/195X
412/195X
413/195X
414/195X
415/195X
416/195X
417/195X
418/195X
419/195X
420/195X
421/195X
422/195X
423/195X
424/195X
425/195X
426/195X
427/195X
428/195X
429

XX
A2
Y39471.

XX
XX

X

| | |
|----|---|
| XX | human; secreted protein; EST; expressed sequence tag; diagnosis; |
| KW | forensic; gene therapy; chromosome mapping; signal peptide; |
| KW | upstream regulatory sequence; cytokine activity; cell proliferation; |
| KW | differentiation; hematopoiesis regulation; tissue growth regulation; |
| KW | reproductive hormone regulation; chemokines; chemokine receptor; hematopoietic; |
| KW | proliferative and differentiation; tumour induction; ps. |

Series (continued)

XX
XX-1559(00)JMI[illegible]

XX

| | | | |
|----|--|---------------|---|
| FT | stem_insp | /rtag d | |
| FT | | /rtag e | |
| FT | | | "nucleotides 120-144 may form a double- |
| FT | | | stranded stem with nucleotides 574-588, |
| FT | | | i.e. the intron forms the loop" |
| XX | | | |
| XX | W9261:449-A. | | |
| XX | 23-JAN 1992. | | |
| XX | 15-JUL 1990; | 90WR-US04586. | |
| XX | 13-JUL-1990; | 90US-0554759. | |
| XX | (GENCO) GEN HOSPITAL CORP. | | |
| XX | Seed B. Aruffo A. Antib. M. | | |
| XX | WPI: 1992-054864/07. | | |
| XX | P-PSDR: R20405. | | |
| XX | New CD53 cell surface antigen and DNA encoding it - for | | |
| XX | immuno-therapy and diagnosis of hematopoietic neoplasms, etc. | | |
| XX | Example 4; Fig 8; 160pp; English. | | |
| XX | A library of recombinants having inserts greater than 0.8kb in | | |
| XX | size was prepared in the pEMV vector (see Q21166) from | | |
| XX | 1 microgram of polyA (plus) RNA isolated from the human | | |
| XX | lymphoblastoid cell line JY. The library was screened for CD7 cDNA | | |
| XX | clones which were isolated using the antibody enrichment method | | |
| XX | (see e.g. Q21164 or Q21165). After the third transfection, COS | | |
| XX | cells were panned with specific anti-CD7 MAb's (i.e. 745 and L245). | | |
| XX | A third supernatant was generated and transformed into E.coli. DNA | | |
| XX | was prepared from the resulting colonies and transfected into COS | | |
| XX | cells. Surface expression of CD7 antigen was detected in 7 of the | | |
| XX | 8 transfected cultures by indirect immunofluorescence. Restriction | | |
| XX | enzyme digest of the corresponding plasmid DNAs revealed two | | |
| XX | species, one contained a 1.2kb insert, and the other a 1.4kb insert, | | |
| XX | (i.e. the 1.2kb sequence plus an intron). | | |
| XX | Sequence 1605 BP; 102 A; 614 C; 479 G; 270 T; 0 other. | | |

```

Query Match      75.2%   Score 15.6;   E0.13;   Length 1665.
Best Local Similarity    89.5%;   Pred. No. 2.1e+02;
Mat. 0.07  17. Cons. Ratio  0. Mismatches  2. Indels  0. Gaps

```

```

QVY      1 aaadaagcgcctggacga 19
          ||||| ||||| |||
DB       1260 AACAAGGCTGTGGTGA 1942

```

```

RESULT 10
114708/c
ID      114708 standard; cDNA; 1665 bp.
XX
AC      114708;
XX
GC      40 GC%-1996. (first entry)
XX
DE      Human cb7 cDNA.
XX
KW      cell surface antigen; cloning; immunoselection; immunotherapy;
KW      therapy; diagnosis; vector; pHB3, GB7, E05, T lymphocyte, ss.
XX
SS      Home sapiens.
XX
Key      Location/Qualifiers
FH exon 1..140
FT //aa a
FT /codon_start 58..60

```

[illegible]

CC (D44, GP5), ICAM, LFA-3, FCRIb, FCRIb, TLISA and Leu8 (see
 CC W86188-62, W86183-52 and W86181-52). cDNA (see V81199) is
 CC specifically claimed. CD7 polypeptide (see W89152), a marker for
 CC the identification of T cell acute lymphoblastic leukaemia, has
 CC been expressed in transfected COS cells.
 XX
 SU Sequence 1665 BP; 302 A; 614 C; 479 G; 270 T; 0 other;

Query Match 75.2%; Score 15.8; DB 20; Length 1665;
 Best Local Similarity 89.5%; Pred. No. 2.1e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 aagaaagcgttgaagaa 19
 IIIIIII IIIII III
 DB 1250 AAGAAAGGATGTCGCGAA 1242

RESULT 14
 V28699/c
 ID V28688 standard; cDNA; 808 BP;
 XX V28688;
 XX
 XX 29-JUL-1998 (first entry)
 XX
 XX Ripeening banana pulp (cDNA clone U-039 SEQ ID NO:46;
 XX Banana, ripening; pulp; Musa acuminata cv. Grand Nain; fruit;
 XX genetic control; tissue sequence); ss.
 XX
 XX Musa acuminata.
 XX W-0811229-A2.
 XX 19-MAR-1998.
 XX
 XX 08-SEP 1997; 97MO-GH02424.
 XX
 XX 25 APR 1997; 97GB-0008366.
 XX 10-SEP-1996; 96GB-0018862.
 XX
 XX (GENE) GENECA LTD.
 XX
 XX Bird CR, Medina-Suarez RDJ, Seymour GR;
 XX WPI: 1998-207389/18.
 XX
 XX Modulation of ripening or tissue sequence in bananas - comprises
 XX use of DNA isolated from ripening banana pulp to produce genetically
 XX modified fruit
 XX
 XX Claim 1: Page 47; 72pp; English.

CC The present sequence represents a cDNA clone isolated from ripening
 CC banana pulp. 57 clones were isolated and are given in V28644 to V28699.
 CC The cDNA clone sequences can be used in a method of modulating ripening
 CC or tissue sequence process in plants of the genus Musa. The method
 CC comprises: (a) inserting into the plant material at least 1 of the 57
 CC sequences (as above); (b) regenerating the plant material; and (c)
 CC selecting from the transformed regenerants, plants with modulated
 CC ripening or tissue sequence characteristics. Also described in the
 CC present invention are: (1) plants, their progeny, seed and material
 CC obtained from the plants, produced by the above method, (2) a vector

CC (D44, GP5), ICAM, LFA-3, FCRIb, FCRIb, TLISA and Leu8 (see
 CC W86188-62, W86183-52 and W86181-52). cDNA (see V81199) is
 CC specifically claimed. CD7 polypeptide (see W89152), a marker for
 CC the identification of T cell acute lymphoblastic leukaemia, has
 CC been expressed in transfected COS cells.
 XX
 SU Sequence 1665 BP; 302 A; 614 C; 479 G; 270 T; 0 other;

Query Match 75.2%; Score 15.8; DB 20; Length 1665;
 Best Local Similarity 89.5%; Pred. No. 2.1e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 aagaaagcgttgaagaa 19
 IIIIIII IIIII III
 DB 1250 AAGAAAGGATGTCGCGAA 1242

RESULT 14
 A50582/c
 ID A50582 standard; cDNA; 1665 BP;
 XX A50582;
 XX
 XX 19-DEC-2000 (first entry)
 XX
 XX Human cell surface antigen CD7 cDNA.
 XX
 XX CD7, cell surface antigen; human immunoselection; panin;
 XX immunosens; diagnosis; immunotherapy; gene therapy;
 XX immune disorder; infection; asthma; immune-complex disease;
 XX amyloidosis; multiple sclerosis; parasitic disease, leukaemia, ss.
 XX
 XX Homo sapiens.
 XX
 XX Key location/qualifiers
 XX CBS 58.1210
 XX /+fad a
 XX /+fete "contains an intron"
 XX 140,569
 XX /+taq b
 XX 1637,1642
 XX /+t34
 XX
 XX 586111093-A.
 XX
 XX 29-AUG-2000.
 XX
 XX 29-DEC-1999; 9905-0191612.
 XX
 XX 01-DEC 1992; 9205-0983647.
 XX 25-FEB 1989; 9805-0160411.
 XX 13-JUL-1989; 8903-0479676.
 XX 24-MAR 1990; 9005-0498809.
 XX 13-JUL-1990; 9005-0553759.
 XX
 XX (GENE) GEN HOSPITAL CORP.
 XX
 XX StamenKovic L, Seed B;
 XX WPI: 2000 586382/55.
 XX P-PSDB: Y96129.
 XX
 XX Isolated nucleic acid molecule encoding the CD7 cell surface antigen,
 XX useful for immunodiagnosis and immunotherapy of immune-mediated,
 XX infections or disorders, e.g. asthma, immune complex disease, parasitic
 XX diseases
 XX
 XX Example 4: FIG 9A-B: 75pp; English
 XX
 XX The present sequence is that of cDNA encoding human cell surface
 XX antigen (cDNA) CD7 (see Y96129). The cDNA was derived from a human
 XX T-cell tumour line HPR-ALL cDNA library using a new protocol for


```

cc sequence was isolated from a rice bacterial artificial chromosome
cc (BAC) library. The RKN polypeptide contains a leucine rich repeat
cc motif, and has a predicted molecular mass of 105.4 kDa. The RKN
cc polypeptide is a putative serine/threonine protein kinase. The RKN
cc polypeptides can be used to provide plants having increased growth
cc and yield. The expression control sequence of RKN can be used to
cc increase expression of a gene product of interest in roots of a plant.
cc It can also be used to produce disease resistance plants.
XX
SQ Sequence 4804 BP; 1021 A; 1476 C; 1346 G; 1061 T; 6 other;
Query Match 74.4%; Score 15.4; DB:J1; Length 4804;
Best Local Similarity 94.1%; Pred. No. 5,400-02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
9Y 2 aqaqaqaqaqaqaqa 18
11111111111111111111
11b 2757 aqaqaqaqaqaqaqa 2775
Search completed: May 1, 2001, 04:59:19
Job time: 7064 sec

```

```

cc function in plants comprising a promoter region which is operably in
cc plant cells, a polynucleotide sequence as defined above, and a
cc transcription termination sequence; and (4) a method of controlling
cc plant pathogenesis comprising the application of an i pathogenic agent to
cc plants of (1).
XX
SQ Sequence 808 BP; 154 A; 243 C; 190 G; 198 T; 44 other;
Query Match 74.4%; Score 15.4; DB:J1; Length 808;
Best Local Similarity 89.9%; Pred. No. 3,100-02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
9Y 4 aaaaqaqaqaqaqa 21
11111111111111111111
11b 427 aaaaqaqaqaqaqaqa 410
Result 15
Z46006
11b Z46006 standard; cDNA; 4804 BP;
XX
Av Z46006;
XX
11b 01 MAY 2001 (first entry)
XX
DE cDNA encoding a receptor-like protein kinase designated RKN.
XX
KW Receptor-like protein kinase; RKN; Serine/threonine protein kinase;
KW plant growth; plant yield; disease resistance plant; SS.
XX
CS 0754 sp.
XX
FH Key location/qualifiers
F1 1042..4418
F1 /std_d
F1 /product "receptor-like protein kinase"
F1 /transcript (pos: 4416..4418, acc: c118)
F1 /note "containing 1 initiation"
F1 1042..4418
F1 /std_b
F1 /number 1
F1 5619..4561
F1 /std_c
F1 /number 1
F1 4063..4418
F1 /std_d
F1 /number 2
XX
FN W200004361 A1.
XX
11b 01 FEB 2001.
XX
F1 21 JUL 1997; 99W00516649.
XX
F1 21 JUL 1997; 99W05120655.
XX
FA (SALK ) SALK INS1 BIOLOGICAL STUDIES.
XX
F1 Z6204 5. 23b 20. 1amb c1;
XX
WPI; 2000 1/1319/16.
11b 18006; 990764.
XX
F1 New isolated plant receptor-like protein kinase, used to produce plants
F1 with increased growth and yield, disease resistant plants and
F1 increasing gene expression in roots.
F1 claim 7; Fig 2; 77pp; English.
XX
F1
F1
F1 The present sequence encodes a receptor-like protein kinase
F1 polypeptide, designated RKN. The protein is a member of the
F1 receptor-like protein kinase (RLK) family. The RKN polynucleotide

```

```

cc sequence was isolated from a rice bacterial artificial chromosome
cc (BAC) library. The RKN polypeptide contains a leucine rich repeat
cc motif, and has a predicted molecular mass of 105.4 kDa. The RKN
cc polypeptide is a putative serine/threonine protein kinase. The RKN
cc polypeptides can be used to provide plants having increased growth
cc and yield. The expression control sequence of RKN can be used to
cc increase expression of a gene product of interest in roots of a plant.
cc It can also be used to produce disease resistance plants.
XX
SQ Sequence 4804 BP; 1021 A; 1476 C; 1346 G; 1061 T; 6 other;
Query Match 74.4%; Score 15.4; DB:J1; Length 4804;
Best Local Similarity 94.1%; Pred. No. 5,400-02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
9Y 2 aqaqaqaqaqaqaqa 18
11111111111111111111
11b 2757 aqaqaqaqaqaqaqa 2775
Search completed: May 1, 2001, 04:59:19
Job time: 7064 sec

```


117: qb_est48:*
 118: qb_est49:*
 119: qb_est50:*
 120: qb_est51:*
 121: qb_est52:*
 122: qb_est53:*
 123: qb_est54:*
 124: qb_est55:*
 125: qb_est56:*
 126: qb_est57:*
 127: qb_est58:*
 128: qb_est59:*
 129: qb_est60:*
 130: qb_est61:*
 131: qb_est62:*
 132: qb_est63:*
 133: qb_est64:*
 134: qb_est65:*
 135: qb_est66:*
 136: qb_est67:*
 137: qb_est68:*
 138: qb_est69:*
 139: qb_est70:*
 140: qb_est71:*
 141: qb_est72:*
 142: qb_est73:*
 143: qb_est74:*
 144: qb_est75:*
 145: qb_est76:*
 146: qb_est77:*
 147: qb_est78:*
 148: qb_est79:*
 149: qb_est80:*
 150: qb_est81:*
 151: qb_est82:*
 152: qb_est83:*
 153: qb_est84:*
 154: qb_est85:*
 155: qb_est86:*
 156: qb_est87:*
 157: qb_est88:*
 158: qb_est89:*
 159: qb_est90:*
 160: qb_est91:*
 161: qb_est92:*
 162: qb_est93:*
 163: qb_est94:*
 164: qb_est95:*
 165: qb_est96:*
 166: qb_est97:*
 167: qb_est98:*
 168: qb_est99:*
 169: qb_est100:*
 170: qb_est101:*
 171: qb_est102:*
 172: qb_est103:*
 173: qb_est104:*
 174: qb_est105:*
 175: qb_est106:*
 176: qb_est107:*
 177: qb_est108:*
 178: qb_est109:*
 179: qb_est110:*
 180: qb_est111:*
 181: qb_est112:*
 182: qb_est113:*
 183: qb_est114:*
 184: qb_est115:*
 185: qb_est116:*
 186: qb_est117:*
 187: qb_est118:*
 188: qb_est119:*
 189: qb_est120:*

190: em_qss_pln1:*
 191: em_qss_pln2:*
 192: em_qss_pro:*
 193: em_qss_rnd1:*
 194: em_qss_rnd2:*
 195: em_qss_rnd3:*
 196: em_qss_rnd4:*
 197: em_qss_rnd5:*
 198: em_qss_vrt1:*
 199: em_qss_vrt2:*
 200: em_qss_vrt3:*
 201: qb_qss1:*
 202: qb_qss2:*
 203: qb_qss3:*
 204: qb_qss4:*
 205: qb_qss5:*
 206: qb_qss6:*
 207: qb_qss7:*
 208: qb_qss8:*
 209: qb_qss9:*
 210: qb_qss10:*
 211: qb_qss11:*
 212: qb_qss12:*
 213: qb_qss13:*
 214: qb_qss14:*
 215: qb_qss15:*
 216: qb_qss16:*
 217: qb_qss17:*
 218: qb_qss18:*
 219: qb_qss19:*
 220: qb_qss20:*
 221: qb_qss21:*
 222: qb_qss22:*
 223: qb_qss23:*
 224: qb_qss24:*
 225: qb_qss25:*
 226: qb_qss26:*
 227: qb_qss27:*
 228: qb_qss28:*
 229: qb_qss29:*
 230: qb_qss30:*
 231: qb_qss31:*
 232: qb_qss32:*
 233: qb_qss33:*
 234: qb_qss34:*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | % | Query | Match | Length | DB | ID | Description |
|--------|-----|-------|-------|-------|-------|----------|-----------|----------|-------------|
| | 1 | 21 | 100.0 | 559 | 165 | HE276194 | 601344452 | HE276194 | 601344452 |
| | 2 | 21 | 100.0 | 585 | 165 | HE390259 | 601285141 | HE390259 | 601285141 |
| | 3 | 21 | 100.0 | 749 | 166 | HE391232 | 601395820 | HE391232 | 601395820 |
| | 4 | 18.4 | 87.6 | 555 | 165 | HE337694 | 804647699 | HE337694 | 804647699 |
| | 5 | 18.4 | 87.6 | 569 | 165 | HE337695 | 804647699 | HE337695 | 804647699 |
| | 6 | 18 | 85.7 | 243 | 115 | AW436749 | 77205 MAR | AW436749 | 77205 MAR |
| | 7 | 18 | 85.7 | 242 | 115 | AW436742 | 77257 MAR | AW436742 | 77257 MAR |
| | 8 | 18 | 85.7 | 508 | 168 | HE710007 | MI-0-AV1- | HE710007 | MI-0-AV1- |
| | 9 | 18 | 85.7 | 511 | 166 | HE387918 | 601282146 | HE387918 | 601282146 |
| | 10 | 18 | 85.7 | 817 | 241 | CNS03K5G | AL247741 | CNS03K5G | AL247741 |
| | 11 | 18 | 85.7 | 964 | 231 | CNS03X1G | AL255057 | CNS03X1G | AL255057 |
| | 12 | 18 | 85.7 | 429 | 158 | HE388562 | 601281855 | HE388562 | 601281855 |
| | 13 | 17.4 | 82.9 | 674 | 166 | CNS0512K | AL349545 | CNS0512K | AL349545 |
| | 14 | 17.4 | 82.9 | 951 | 232 | CNS00198 | AL142671 | CNS00198 | AL142671 |
| | 15 | 17 | 81.0 | 1101 | 229 | CNS02DHQ | HE575394 | CNS02DHQ | HE575394 |
| | 16 | 17 | 81.0 | 66 | 229 | HE575394 | HE575394 | HE575394 | HE575394 |
| | 17 | 15.8 | 79.0 | 66 | 229 | HE575394 | HE575394 | HE575394 | HE575394 |
| | 18 | 16.8 | 80.0 | 255 | 160 | HE575394 | HE575394 | HE575394 | HE575394 |

```

19 16.38 80.0 36.7 137 BE950829
20 16.38 80.0 4.4 141 BE950829
21 16.38 80.0 4.7 241 A1716302
22 16.38 80.0 5.0 146 BE293367
23 16.38 80.0 6.1 145 BE253944
24 16.38 80.0 6.4 227 BE253944
25 16.38 80.0 6.5 137 BE650433
26 16.38 80.0 7.7 13 A0920189
27 16.38 80.0 8.0 146 BE309894
28 16.38 80.0 9.4 145 BE293367
29 16.38 80.0 2.1 157 BE439471
30 16.38 80.0 2.4 14 A0967180
31 16.38 80.0 2.8 1 A096813
32 16.38 80.0 2.7 13 A095591
33 16.38 80.0 3.5 10 A0572826
34 16.38 80.0 3.2 24 A1716302
35 16.38 80.0 3.4 159 BE0866790
36 16.38 80.0 3.4 164 BE0866790
37 16.38 80.0 3.6 158 BE0866790
38 16.38 80.0 3.6 158 BE0866790
39 16.38 80.0 3.6 158 BE0866790
40 16.38 80.0 3.6 158 BE0866790
41 16.38 80.0 3.6 158 BE0866790
42 16.38 80.0 3.6 158 BE0866790
43 16.38 80.0 3.6 158 BE0866790
44 16.38 80.0 3.6 158 BE0866790
45 16.38 80.0 3.6 158 BE0866790

```

ALIGNMENTS

```

RESULT 1
LOCUS BE276194 559 bp. mRNA 14-JUL-2000
DEFINITION GenBank: BE276194.1 Homo sapiens cDNA clone IMAGE:4059630.
VERSION BE276194.1 GI:1911157
KEYWORDS Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 559)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Published (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: AFCE
cDNA Library Prepared by: Ling Hong/Rubin Laboratory
cDNA Library Adapted by: The I.M.A.G.E. Consortium (ILRG)
DNA Sequencing by: Lucyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILRG at:
http://image.fhc.org
Plate: L10M257 Row: m Column: 08
High quality sequence step: 559.
Location/Qualifiers
1..559

```

```

FEATURES
Source
1..559
Organization "Homo sapiens"
Zelone "IMAGE:4059630"
Zelone "LIB "MIM 367220"
Zelone type "epidemiologic melanoma cell line"
Zelone host "DH10b (phage resistant)"
Zelone "Clon. Vector: pTZ19, Clon. Site: 2"
Zelone cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XbaI sites using the following 5'
adaptor: GAGGAGG(5'). Library constructed by Ling Hong
in the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 106 a 171 c 232 g 76 t
ORIGIN
1..559

```

```

GenBank
Host Local Similarity 100.0% Score 21; 108 166; Length 586;
Matches 21; Conserved 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aadaaaagagctatagaatg 21
|||||
106 427 AGGGAAGGCGTGGGAGGAG 447

RESULT 4
LOCUS BE391242
DEFINITION GenBank: BE391242.1 Homo sapiens cDNA clone IMAGE:4059630.
VERSION BE391242.1 GI:1911157
KEYWORDS Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 559)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Published (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: AFCE
cDNA Library Prepared by: Ling Hong/Rubin Laboratory
cDNA Library Adapted by: The I.M.A.G.E. Consortium (ILRG)
DNA Sequencing by: Lucyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILRG at:
http://image.fhc.org
Plate: L10M257 Row: m Column: 08
High quality sequence step: 559.
Location/Qualifiers
1..559

```

```

FEATURES
Source
1..559
Organization "Homo sapiens"
Zelone "IMAGE:4059630"
Zelone "LIB "MIM 367220"
Zelone type "epidemiologic melanoma"
Zelone host "DH10b (phage resistant)"
Zelone "Clon. Vector: pTZ19, Clon. Site: 2"
Zelone cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XbaI sites using the following 5'
adaptor: GAGGAGG(5'). Size selected ~500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 106 a 171 c 232 g 76 t
ORIGIN
1..559

```

```

California, Berkeley) using ZAP cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 98 a 161 c 212 g 86 t
ORIGIN
1..586
GenBank
Host Local Similarity 100.0% Score 21; 108 166; Length 586;
Matches 21; Conserved 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aadaaaagagctatagaatg 21
|||||
106 427 AGGGAAGGCGTGGGAGGAG 447

RESULT 2
LOCUS BE390259 586 bp. mRNA 21-JUL-2000
DEFINITION GenBank: BE390259.1 Homo sapiens cDNA clone IMAGE:4059630.
VERSION BE390259.1 GI:1911157
KEYWORDS Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 586)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Published (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: AFCE
cDNA Library Prepared by: Ling Hong/Rubin Laboratory
cDNA Library Adapted by: The I.M.A.G.E. Consortium (ILRG)
DNA Sequencing by: Lucyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILRG at:
http://image.fhc.org
Plate: L10M257 Row: m Column: 08
High quality sequence step: 586.
Location/Qualifiers
1..586

```

```

Organization "Homo sapiens"
Zelone "IMAGE:4059630"
Zelone "LIB "MIM 367220"
Zelone type "epidemiologic melanoma cell line"
Zelone host "DH10b (phage resistant)"
Zelone "Clon. Vector: pTZ19, Clon. Site: 2"
Zelone cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XbaI sites using the following 5'
adaptor: GAGGAGG(5'). Library constructed by Ling Hong
in the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 106 a 171 c 232 g 76 t
ORIGIN
1..586

```

```

GenBank
Host Local Similarity 100.0% Score 21; 108 166; Length 586;
Matches 21; Conserved 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aadaaaagagctatagaatg 21
|||||
106 427 AGGGAAGGCGTGGGAGGAG 447

RESULT 4
LOCUS BE391242
DEFINITION GenBank: BE391242.1 Homo sapiens cDNA clone IMAGE:4059630.
VERSION BE391242.1 GI:1911157
KEYWORDS Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 586)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Published (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: AFCE
cDNA Library Prepared by: Ling Hong/Rubin Laboratory
cDNA Library Adapted by: The I.M.A.G.E. Consortium (ILRG)
DNA Sequencing by: Lucyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILRG at:
http://image.fhc.org
Plate: L10M257 Row: m Column: 08
High quality sequence step: 586.
Location/Qualifiers
1..586

```

```

Organization "Homo sapiens"
Zelone "IMAGE:4059630"
Zelone "LIB "MIM 367220"
Zelone type "epidemiologic melanoma cell line"
Zelone host "DH10b (phage resistant)"
Zelone "Clon. Vector: pTZ19, Clon. Site: 2"
Zelone cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XbaI sites using the following 5'
adaptor: GAGGAGG(5'). Library constructed by Ling Hong
in the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 106 a 171 c 232 g 76 t
ORIGIN
1..586

```

```

GenBank
Host Local Similarity 100.0% Score 21; 108 166; Length 586;
Matches 21; Conserved 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aadaaaagagctatagaatg 21
|||||
106 427 AGGGAAGGCGTGGGAGGAG 447

RESULT 4
LOCUS BE391242
DEFINITION GenBank: BE391242.1 Homo sapiens cDNA clone IMAGE:4059630.
VERSION BE391242.1 GI:1911157
KEYWORDS Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 586)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Published (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: AFCE
cDNA Library Prepared by: Ling Hong/Rubin Laboratory
cDNA Library Adapted by: The I.M.A.G.E. Consortium (ILRG)
DNA Sequencing by: Lucyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILRG at:
http://image.fhc.org
Plate: L10M257 Row: m Column: 08
High quality sequence step: 586.
Location/Qualifiers
1..586

```

LOCUS BE391242 749 bp mRNA EST 21-JUL-2000
 DEFINITION 601265820P1 NH_MGC_44 Homo sapiens cDNA clone IMAGE:3607708 5',
 mRNA sequence.
 ACCESSION BE391242
 VERSION BE391242.1 GI:9336597
 KEYWORDS EST.
 SOURCE Human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 739)
 AUTHORS NIH MGC http://www.ncbi.nlm.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Tel: (401) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILLN at:
 http://image.llnl.gov/
 plates: L10M259 row: q column: 05
 High quality sequence stop: 631.
 location/Qualifiers
 1..739
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:3607708"
 /clone_lib="NH_MGC_44"
 /accession_type="cDNA; cDNA, adenoviral; inema cell line"
 /lab_host="pR10B (phage-resistant)"
 /note="Organ: uterus; Vector: pOT87; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGAAGAGG(s). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 121 a 213 c 283 g 122 t
 ORIGIN
 Query Match 100.0% Score 21; DB 166; Length 739;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 aaaaauuuccttgcacatg 21
 ++++++|||||
 DB 426 AGGAAAGAGCTGGAGGATG 446
 RESULT 4
 BE337694/5 555 bp mRNA EST 14-JUL-2000
 LOCUS 894047609.y1 C. reinhardtii CC-1690, normalized, Lambda Zap 11
 DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.
 ACCESSION BE337695
 VERSION BE337695.1 GI:9210780
 KEYWORDS EST.
 SOURCE Chlamydomonas reinhardtii.
 ORGANISM Chlamydomonas reinhardtii
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadales; Chlamydomonas.
 REFERENCE 1 (bases 1 to 555)
 AUTHORS Grossman, A., Davies, J., Pederspiel, N., Harris, E., Lefebvre, P.,
 McMorris, J.P., Silliman, C., Stern, D., and Szarycki, R.
 TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
 Unicellular System for Analyzing Gene Function and Regulation in
 Vascular Plants; project phase 2
 JOURNAL Unpublished (2000)
 COMMENT Contact: Elizabeth H. Harris
 DCMR Box 91000
 Duke University
 Durham, NC 27708-1000, USA
 Tel: 919 613 8164
 Fax: 919 613 8177
 Email: chlamy@duke.edu.
 location/Qualifiers
 1..555
 /organism="Chlamydomonas reinhardtii"
 /strain="CC-1690 wild type mt, 21qt"
 /db_xref="taxon:3055"
 /clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
 11"
 Query Match 87.6% Score 18.4; DB 165; Length 555;
 Best Local Similarity 95.0%; Pred. No. 4e+02; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 1;
 QY 2 ggaagaaacatgcacacatg 21
 ++++++|||||
 DB 460 GGGAAAGAGCTGGAGGATG 441
 RESULT 5
 BE337695 569 bp mRNA EST 14-JUL-2000
 LOCUS 894047609.y1 C. reinhardtii CC-1690, normalized, Lambda Zap 11
 DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.
 ACCESSION BE337695
 VERSION BE337695.1 GI:9210780
 KEYWORDS EST.
 SOURCE Chlamydomonas reinhardtii.
 ORGANISM Chlamydomonas reinhardtii
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadales; Chlamydomonas.
 REFERENCE 1 (bases 1 to 569)
 AUTHORS Grossman, A., Davies, J., Pederspiel, N., Harris, E., Lefebvre, P.,
 McMorris, J.P., Silliman, C., Stern, D., and Szarycki, R.
 TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
 Unicellular System for Analyzing Gene Function and Regulation in
 Vascular Plants; project phase 2
 JOURNAL Unpublished (2000)
 COMMENT Contact: Elizabeth H. Harris
 DCMR Box 91000
 Duke University
 Durham, NC 27708-1000, USA
 Tel: 919 613 8164
 Fax: 919 613 8177
 Email: chlamy@duke.edu.
 location/Qualifiers
 1..569
 /organism="Chlamydomonas reinhardtii"
 /strain="CC-1690 wild type mt, 21qt"
 /db_xref="taxon:3055"
 /clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
 11"

COMMENT Contact: Elizabeth H. Harris
 DCMR Box 91000
 Duke University
 Durham, NC 27708-1000, USA
 Tel: 919 613 8164
 Fax: 919 613 8177
 Email: chlamy@duke.edu.
 location/Qualifiers
 1..555
 /organism="Chlamydomonas reinhardtii"
 /strain="CC-1690 wild type mt, 21qt"
 /db_xref="taxon:3055"
 /clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
 11"
 Znote: "Vector: p Bluescript II SK+; Site_1: EcoRI; Site_2:
 XhoI; This library, constructed by John Davies and Jeffrey
 McMorris, combines cDNAs from CC-1690 cells grown to
 mid-log phase in TAP (acetate-containing) medium in the
 light, TAP medium in the dark, HS (minimal) medium in
 ambient levels of CO2 and HS medium bubbled with 5% CO2.
 PolyA mRNA was purified from each sample, pooled and cDNA
 synthesized. The cDNA was directionally cloned into Lambda
 ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
 Bluescript II SK+ plasmids were excised from the Lambda
 ZAP clones by superinfection with ExAssist (Stratagene)
 phage. The library was normalized using method 4 described
 in Bonaldo et al (1996) Genome Research 6: 791-806."
 BASE COUNT 127 a 183 c 141 g 114 t
 ORIGIN
 Query Match 87.6% Score 18.4; DB 165; Length 555;
 Best Local Similarity 95.0%; Pred. No. 4e+02; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 1;
 QY 2 ggaagaaacatgcacacatg 21
 ++++++|||||
 DB 460 GGGAAAGAGCTGGAGGATG 441

RESULT 5
 BE337695 569 bp mRNA EST 14-JUL-2000
 LOCUS 894047609.y1 C. reinhardtii CC-1690, normalized, Lambda Zap 11
 DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.
 ACCESSION BE337695
 VERSION BE337695.1 GI:9210780
 KEYWORDS EST.
 SOURCE Chlamydomonas reinhardtii.
 ORGANISM Chlamydomonas reinhardtii
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadales; Chlamydomonas.
 REFERENCE 1 (bases 1 to 569)
 AUTHORS Grossman, A., Davies, J., Pederspiel, N., Harris, E., Lefebvre, P.,
 McMorris, J.P., Silliman, C., Stern, D., and Szarycki, R.
 TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
 Unicellular System for Analyzing Gene Function and Regulation in
 Vascular Plants; project phase 2
 JOURNAL Unpublished (2000)
 COMMENT Contact: Elizabeth H. Harris
 DCMR Box 91000
 Duke University
 Durham, NC 27708-1000, USA
 Tel: 919 613 8164
 Fax: 919 613 8177
 Email: chlamy@duke.edu.
 location/Qualifiers
 1..569
 /organism="Chlamydomonas reinhardtii"
 /strain="CC-1690 wild type mt, 21qt"
 /db_xref="taxon:3055"
 /clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
 11"

Gene "Vector" plasmid; 11 SP, Site_1: XbaI; Site_2: XbaI; this library constructed by 32n larvae and after mid-lar phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, BS (minimal) medium in ambient levels of 0.2 and 0.5 medium bubbled with 5% O₂. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda ZAP II (Stratagene) in the EcoRI (5') and XbaI (3') sites. ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldi et al (1996) Genome Research 6: 791-806."

BASE COUNT 102 4 140 3 233 4 123 1 1 others

Query Match 87.6%; Score 18.4; DB 165; Length 569;
Best Local Similarity 95.0%; Pred. No. 4002;
Matches 18; Conservation 0; Mismatches 1; Indels 0; Gaps 0;

27 2 qqaaadgaagctatgaagca 21

|||||1111111111111111

DB 463 GGGAGAGGCTGGGGGA 482

RESULT 6

AW436749

DEFINITION 77296 MARC 211G Sus scrofa cDNA 5', mRNA sequence.

ACCESSION AW436749

VERSION AW436749.1 GI:6972955

KEYWORDS EST

SOURCE pig

ORGANISM

Sus scrofa

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

1 (bases 1 to 243)

Fahrenkrug, S.C., Froking, B.A., Rehner, G.A., Smith, T.F.L., Casas, E.,

Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laqureid, W.W.

and Keele, J.W.

Design and use of two pooled tissue normalized cDNA libraries for

EST discovery in swine

Unpublished (2000)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@mail.mars.usda.gov

Single pass sequencing, bases called and trimmed with phred

20,980904.c, Vector identified by cross_match with the -minscore 20

and -mismatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACAT

BACKWARD: GTTTCGACGACGACG

Plate: 32 row: 1 column: 17

Seq primer: ATTAGGTGACACTATAG.

Location/Qualifiers

1..243

Zenonism "Sus scrofa"

Zb_xref "x000924"

Zclone_lib "MARC 211G"

Zissue_type "pooled"

Zlab_host "BHL08"

Znote "Vector: pCMV Spacetr; Site_1: XbaI; Site_2: XbaI;

Library made from pooled tissue from testis, ovary,

endometrium, hypothalamus, pituitary, and placenta."

39 4 75 3 77 4 42 1

BASE COUNT 102 4 140 3 233 4 123 1 1 others

Query Match 85.7%; Score 19; DB 115; Length 294;

Best Local Similarity 100.0%; Pred. No. 57002;
Matches 18; Conservation 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 qqaaadgaagctatgaagca 19

|||||1111111111111111

DB 110 GGGAGAGGCTGGGGGA 127

RESULT 7

AW436782

DEFINITION 77257 MARC 211G Sus scrofa cDNA 5', mRNA sequence.

ACCESSION AW436782

VERSION AW436782.1 GI:6972088

KEYWORDS EST

SOURCE pig

ORGANISM

Sus scrofa

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

1 (bases 1 to 242)

Fahrenkrug, S.C., Froking, B.A., Rehner, G.A., Smith, T.F.L., Casas, E.,

Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laqureid, W.W.

and Keele, J.W.

Design and use of two pooled tissue normalized cDNA libraries for

EST discovery in swine

Unpublished (2000)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@mail.mars.usda.gov

Single pass sequencing, bases called and trimmed with phred

20,980904.c, Vector identified by cross_match with the -minscore 20

and -mismatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACAT

BACKWARD: GTTTCGACGACGACG

Plate: 32 row: 1 column: 20

Seq primer: ATTAGGTGACACTATAG.

Location/Qualifiers

1..242

Zenonism "Sus scrofa"

Zb_xref "x000924"

Zclone_lib "MARC 211G"

Zissue_type "pooled"

Zlab_host "BHL08"

Znote "Vector: pCMV Spacetr; Site_1: XbaI; Site_2: XbaI;

Library made from pooled tissue from testis, ovary,

endometrium, hypothalamus, pituitary, and placenta."

39 4 78 3 81 4 44 1

BASE COUNT 102 4 140 3 233 4 123 1 1 others

Query Match 85.7%; Score 18; DB 115; Length 242;

Best Local Similarity 100.0%; Pred. No. 57002;

Matches 18; Conservation 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 qqaaadgaagctatgaagca 19

|||||1111111111111111

DB 110 GGGAGAGGCTGGGGGA 127

RESULT 8

BE234043

DEFINITION 139049 MARC 111G Sus scrofa cDNA 5', mRNA sequence.

ACCESSION BE234043

VERSION BE234043.1 GI:9017761

KEYWORDS EST

SOURCE pig

ORGANISM

Sus scrofa

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;


```

Zelone "IMAGS6030804"
Zelone Lib "NHL_M2_44"
Zelone Type "Random Prim, anti-mouse/human cell line"
Zelone Host "pH10B (phage resistant)"
Zelone Vector "Vector: pTZ19; Site 1: XhoI; Site 2:
Zelone "cDNA made by oligo-dT priming, directional
cloned into EcoRI/XbaI sites using the following 5'
adaptor: GGAAGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP cDNA synthesis kit
(Spartanone) and Superscript II RT (Life Technologies)."
BASE COUNT      164  147  213  743
CGCTTIN

Query Match      85.7%  Score 19; DB 166; Length 511;
Best local Similarity 107.06; Freq. No. 5,9602;
Matches 18; Conservative 0; Mismatches 0; Gaps 0;

ID: 448 GGAAGAGGCTGAGGAGA 441
|||||
2 appaaggaactaggaaga 19
|||||

RESULT 12
LOCUS      CNS03X15/6
DEFINITION Tetraodon nitroviridis genome survey sequence, p00-ori end of clone
064804 of library 6 from Tetraodon nitroviridis, genome survey
sequence.
ACCESSION   AL265057
VERSION     AL265057.1 GI:7086777
KEYWORDS   GSS; genome survey sequence.
SOURCE      Tetraodon nitroviridis.
ORGANISM    Tetraodon nitroviridis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Eurypterygii; Clupeosquamata; Acanthomorpha; Eucanthomorpha;
Holacanthopterygii; Acanthopterygii; Perciformes;
Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.
1 (bases 1 to 964)
Kornfield, R., Jullienne, Basilva, C., Fizames, C., Fisher, C.,
Bennet, A., Frazmes, C., Quetier, P., Saurin, W., Bernot, A. and
Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish tetraodon nitroviridis
Unpublished
2 (bases 1 to 964)
Kornfield, R., Jullienne, Basilva, C., Fizames, C., Bernot, A.,
Bennet, A., Frazmes, C., Wincker, P., Brattier, P., Quetier, P.,
Saurin, W., and Weissenbach, J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nitroviridis DNA sequence
Unpublished
3 (bases 1 to 964)
Genoscope.
Direct Submission
Submitted (12 APR 2000) to the EMBL/Genbank/DBCLS databases
This sequence is a single read and was generated as part of a large
scale clone and sequencing project of the Tetraodon nitroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/53/seq/2001/tet
FEATUSES
SOURCE      Kornfield, R., Jullienne, Basilva, C., Fizames, C.,
Bennet, A., Frazmes, C., Quetier, P., Saurin, W., Bernot, A. and
Weissenbach, J.
Tetraodon nitroviridis
1, 964
Zelone "IMAGS6030804"
Zelone Lib "NHL_M2_44"
Zelone Type "Random Prim, anti-mouse/human cell line"
Zelone Host "pH10B (phage resistant)"
Zelone Vector "Vector: pTZ19; Site 1: XhoI; Site 2:
Zelone "cDNA made by oligo-dT priming, directional
cloned into EcoRI/XbaI sites using the following 5'
adaptor: GGAAGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP cDNA synthesis kit
(Spartanone) and Superscript II RT (Life Technologies)."
BASE COUNT      164  147  213  743
CGCTTIN

Query Match      85.7%  Score 19; DB 166; Length 511;
Best local Similarity 107.06; Freq. No. 5,9602;
Matches 18; Conservative 0; Mismatches 0; Gaps 0;

ID: 448 GGAAGAGGCTGAGGAGA 441
|||||
2 appaaggaactaggaaga 19
|||||

RESULT 13
LOCUS      W64762/4
DEFINITION m897c08.11 Scary genome end; y6 RM8911.1 14.5 Mb Australian cDNA
W64762/4
BASE COUNT      164  147  213  743
CGCTTIN

```


Query Match 100.0%; Score 21; DB 1; Length 1002;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaaaagagcctagcagatg 21
IIIIIIIIIIIIIIIIIIIIII

DB 488 AGGAAAGCGCTGGGCGATG 508

RESULT 2

US-09-345-409-1
Sequence 1: Application US/09-345409
Patent No. 6121029
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Lison, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyt, Devon
APPLICANT: Goettlich, Joern
TITLE OF INVENTION: GINS FOR THE RE-SYNTHESIS OF PROTHROMBINS
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/345,409
CURRENT FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 40
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 1
LENGTH: 68750
TYPE: DNA
ORGANISM: Sorandium cellulosum
US 09-345-409-1

Query Match 73.4%; Score 15.4; DB 4; Length 68750;
Best Local Similarity 94.1%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 aaaaagagcctagcag 17
IIIIIIIIIIIIIIIIIIIIII

DB 41747 aaaaagagcctagcag 31764

RESULT 3

US 09-346-804-406
Sequence 45: Application US/09-346804
Patent No. 6150857
GENERAL INFORMATION:
APPLICANT: Ralph, David
APPLICANT: Wu, Scott
APPLICANT: O'Hara, Mark S.
APPLICANT: Weller, Robert
TITLE OF INVENTION: DIAGNOSIS OF DISEASE STATE USING mRNA
TITLE OF INVENTION: PROFILES IN PLETHRAL LIQUOR
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkso
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER RELEASABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/346,804
FILING DATE: Concurrently Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/041,576
FILING DATE: 24-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Nakashima, Richard A.
REGISTRATION NUMBER: P 42,024
REFERENCE/BOOKLET NUMBER: 09001014
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 183 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US 09-046-894-49

Query Match 72.4%; Score 15.2; DB 4; Length 184;
Best Local Similarity 85.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 aggaagagcctagcagatg 20
IIIIIIIIIIIIIIIIIIIIII

DB 110 AGGAAAGCGCTGGGCGATG 91

RESULT 4

US-08-485-229-3
Sequence 3: Application US/08-485229
Patent No. 5605690
GENERAL INFORMATION:
APPLICANT: Jacobs, Cindy A.
APPLICANT: Smith, Craig A.
TITLE OF INVENTION: Method of Treating TNF-independent
TITLE OF INVENTION: Inflammation Using Tumor Necrosis Factor Antagonists
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101
COMPUTER RELEASABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,229
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/946,246
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wight, Christopher L.
REGISTRATION NUMBER: 41,680
REFERENCE/BOOKLET NUMBER: 256;
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 587-0606
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1557 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:


```

QY 1 aqaaagagcgcctagcgcq 18
+| | | | | | | | | | | | | | | |
db 19 AGAAGAGGCGCTGGGCTG 46

RESULT 7
US-08-650-000-1
: Sequence 1, Application US/08650000
: Patent No. 5395760
: GENERAL INFORMATION:
: APPLICANT: Smith, Craig A.
: APPLICANT: Goodwin, Raymond G.
: APPLICANT: Beckmann, M. P.
: TITLE OF INVENTION: Tumor Necrosis Factor Receptors
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Immunex Corporation
: STREET: 51 University Street
: CITY: Seattle
: STATE: Washington
: COUNTRY: U.S.A.
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Pafin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/650,000
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/468,453
: FILING DATE:
: APPLICATION NUMBER: US/08/000,766
: FILING DATE:
: APPLICATION NUMBER: US 404,241
: FILING DATE: 05-SEP-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 405,370
: FILING DATE: 11-SEP-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 524,635
: FILING DATE: 10-MAY-1990
: ALTERNATE INFORMATION:
: NAME: Wirth, Christopher L.
: REGISTRATION NUMBER: 41,680
: REFERENCE CYCLET NUMBER: 2501 D
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 587-0440
: TELEFAX: (206) 243 0644
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1641 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: Single
: TOPOLOGY: Linear
: MOLECULE TYPE: cDNA to mRNA
: HYDROTICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Homo sapiens
: CELL TYPE: Fibroblast
: CELL LINE: WI-26 VA4
: IMMEDIATE SOURCE:
: LIBRARY: WI-26 VA4
: CLONE: 1
: FEATURE:
: NAME/KEY: CDS

```

```

: LOCATION: 88...1473
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 154...1470
: FEATURE:
: NAME/KEY: sig_peptide
: LOCATION: 88...153
: PUBLICATION INFORMATION:
: AUTHORS: Smith, Craig A.
: AUTHORS: Davis, Jorri
: AUTHORS: Anderson, Dirk
: AUTHORS: Solam, Lisabeth
: AUTHORS: Beckmann, M. P.
: AUTHORS: Jerzy, Rita
: AUTHORS: Lower, Steven K.
: AUTHORS: Cosman, David
: AUTHORS: Goodwin, Raymond G.
: TITLE: A Receptor for Tumor Necrosis Factor Defines
: TITLE: an Unusual Family of Cellular and Viral Proteins
: JOURNAL: Science
: VOLUME: 248
: PAGES: 1019-1023
: DATE: 25-MAY-1990
: US-08-650-000-1

Query Match 70.5%; Score 14.8; DB 2; Length 1641;
Best local Similarity 88.9%; Pred. No. 20:02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 aqaaagagcgcctagcgcq 18
+| | | | | | | | | | | | | | | |
db 19 AGAAGAGGCGCTGGGCTG 46

RESULT 8
US-08-722-001-1
: Patent No. 5395760
: APPLICANT: SMITH, CRAIG A., JERZY, RITA, KAYM, RD G., BECKMANN,
: M. PATRICIA
: TITLE OF INVENTION: RNA ENCODING TUMOR NECROSIS FACTOR 4 AND
: RECEPTORS
: NUMBER OF SEQUENCES: 17
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/533,635
: FILING DATE: 10-MAY-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 421,417
: FILING DATE: 13-OCT-1989
: APPLICATION NUMBER: 405,370
: FILING DATE: 11-SEP-1989
: APPLICATION NUMBER: 404,241
: FILING DATE: 05-SEP-1989
: SEQ ID NO: 1
: LENGTH: 1641
: US-08-722-001-1

Query Match 70.5%; Score 14.8; DB 6; Length 1641;
Best local Similarity 88.9%; Pred. No. 20:02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 aqaaagagcgcctagcgcq 18
+| | | | | | | | | | | | | | | |
db 19 AGAAGAGGCGCTGGGCTG 46

RESULT 9
US-08-722-001-29/c
: Sequence 29, Application US/08722001
: Patent No. 5760054
: GENERAL INFORMATION:
: APPLICANT: Thompson, Wayne J.

```



```

FI      exon      /number 5
FI      /tag      4897..4996
FI      /tag      n
FI      /number 7
FI      /note- "Although the specification states that the
FI      exon/intron boundary is between position 4996 and
FI      4997 it is believed to be at position 4995 as a
FI      slice at position 4996-4997 alters the
FI      translation of the protein represented in W36509"
FI      intron     4997..5305
FI      /tag      o
FI      /number 7
FI      /tag      5306..5414
FI      /tag      p
FI      /number 8
FI      /tag      5415..5498
FI      /tag      q
FI      /number 8
FI      /tag      5499..5658
FI      /tag      r
FI      /number 9
FI      /tag      5659..5733
FI      /tag      s
FI      /number 9
FI      /tag      5734..5852
FI      /tag      t
FI      /number 10
FI      /tag      5853..6153
FI      /tag      u
FI      /number 10
FI      /tag      6154..6318
FI      /tag      v
FI      /number 11
FI      /tag      6319..6394
FI      /tag      w
FI      /number 11
FI      /tag      6395..6509
FI      /tag      x
FI      /number 12
FI      /tag      6510..7082
FI      /tag      y
FI      /number 12
FI      /tag      7083..7698
FI      /tag      z
FI      /number 13
FI      /note "Although the specification states that this
FI      exon/intron boundary is between position 7698 and
FI      7699 there is a putative intron within this
FI      exon at position 7236 to 7284. The sliced
FI      out of this intron allows a reading frame which
FI      translates to the protein represented in W36509"
FI      intron     7699..8943
FI      /tag      aa
FI      /number 14
FI      /tag      8944..9062
FI      /tag      ab
FI      /number 14
FI      /tag      9063..9427
FI      /tag      ac
FI      /number 14
FI      /tag      9428..9483
FI      /tag      ad
FI      /number 15
FI      /tag      9484..9955
FI      /tag      ae
FI      /number 15
FI      /tag      9956..10100
FI      /tag      af
FI      /number 16
FI      /tag      10101..10470
FI      /tag      ag
FI      /number 16
FI      /tag      10471..10644

```

```

FI      /tag      ah
FI      /number 17
FI      /note "Although the specification states that this
FI      exon/intron boundary is between position
FI      10644 and 10645 it is believed to be at
FI      position 10650 to 10651 as a slice at
FI      position 10644 to 10645 results in the loss of
FI      the Val and Gly residues at positions 849 and
FI      850 of the protein W36509"
FI      intron     10645..10742
FI      /tag      ai
FI      /number 17
FI      /tag      10743..10826
FI      /tag      aj
FI      /number 18
FI      /tag      10827..11286
FI      /tag      ak
FI      /number 18
FI      /tag      11287..11448
FI      /tag      al
FI      /number 19
FI      /tag      11449..11528
FI      /tag      am
FI      /number 19
FI      /tag      11529..11744
FI      /tag      ap
FI      /number 20
FI      /tag      11745..11974
FI      /tag      aq
FI      /number 20
FI      /tag      11974..12090
FI      /tag      ar
FI      /number 21

```

W09740855-A1.

06 NOV-1997.

01 OCT-1996; 96W0-0815769.

29 APR-1996; 96JC 3016482.

(U) J. J. UNIV JOHN HOPKINS SCHOL MEDICINE.

Dietz BC;

W11; 1997-549494/50.

P-PS106; W36509.

FI New isolated regulators of non sense mediated RNA decay - used to
FI develop products for the study, diagnosis and therapy of disorders
FI such as Marfan Syndrome, accelerated ageing and cancers

Claim 1; Fig 4; 79pp; English.

CC this sequence encodes the murine RENT1 protein which regulates
CC telomerase mediated RNA decay (EMBO). The RENT1 (product of non-sense
CC transcripts) protein and other products can be used in the study,
CC diagnosis and therapy of disorders involving NMD such as Marfan
CC Syndrome, accelerated aging or various cancers.

CC Sequence 13146 BP; 2863 A; 3590 G; 671 G; 3222 T; 0 other;

Query Match

Best Local Similarity 47.9%; Score 24; DB 18; Length 13146;

Matches 29; Conservative 0; Mismatches 10; Gaps 0;

Query Match 100.0%; Score 48; DB 1; Length 1002;
 Best Local Similarity 100.0%; Pred. No. 3,7e-19;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aaactaaagatctcaaacaccccaaacacaaagttctcaaacacaa 48
 |||||
 DB 554 AAAATACAGATCTCAAAACCAACCAACCAAGGTTCAGAACAGAGA 601

RESULT 2
 US-08 791-495-1
 : Sequence 1, Application US/08791495
 : Patent No. 5811519
 : GENERAL INFORMATION:
 : APPLICANT: Leth, Ronald
 : APPLICANT: Leach, Sophie
 : APPLICANT: De Smet, Charles
 : APPLICANT: Capelgite, Patrick
 : APPLICANT: Boon-Fallier, Thierry
 : TITLE OF INVENTION: LI-1 TUMOR SPECIFIC GENES
 : NUMBER OF SEQUENCES: 14
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Wolf, Greenfield & Sacke, P.C.
 : STREET: 600 Atlantic Avenue
 : CITY: Boston
 : STATE: MA
 : COUNTRY: USA
 : ZIP: 02210

: COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: 08/088,791.495
 : FILING DATE:
 : CLASSIFICATION: 4.35
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Van Amsterdam, John R.
 : REGISTRATION NUMBER: 40,212,400,000
 : TELEPHONE: 617-720-2441
 : TELEFAX: 617-720-2441
 : INFORMATION FOR SEQ ID NO: 1:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 217 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: double
 : TOPOLOGY: linear
 : MOLECULE TYPE: cDNA
 : HYPOTHETICAL: NO
 : ANTI-SENSE: NO
 : FRAGMENT TYPE: internal
 : ORIGINAL SOURCE:
 : ORGANISM: Homo sapiens
 : US-08 791-495-1

Query Match 81.2%; Score 39; DB 1; Length 217;
 Best Local Similarity 100.0%; Pred. No. 8.1e-07;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 gatctcaaacaccccaaacacaaagttctcaaacacaa 48
 |||||
 DB 1 GATCTCAGAAACCAACCAACCAAGGTTCAGAACAGAGA 49

RESULT 3
 US-08 724-454D-3/c
 : Sequence 3, Application US/080724454D

: Patent No. 594119
 : GENERAL INFORMATION:
 : APPLICANT: Dietz, Harry C.
 : TITLE OF INVENTION: MAMMALIAN REGULATOR OF
 : TITLE OF INVENTION: NONSENSE MEDIATED RNA DEGRAD
 : NUMBER OF SEQUENCES: 29
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Fish & Richardson, P.C.
 : STREET: 4225 Executive Square, Suite 1400
 : CITY: La Jolla
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 92037
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: Windows95
 : SOFTWARE: FastSeq for Windows Version 2.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: 08/033,372.4, 354D
 : FILING DATE: 01-OCT-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 60/016,482
 : FILING DATE: 29-APR-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Hille, Lisa A.
 : REGISTRATION NUMBER: 48,347
 : REFERENCE/AGENT NUMBER: 07265,0900001
 : TELEPHONE/AGENT INFORMATION:
 : TELEPHONE: 619-678-5070
 : TELEFAX: 619-678-5099
 : INFORMATION FOR SEQ ID NO: 3:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 13146 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: genomic DNA
 : US-08-724-354D-3

Query Match 47.9%; Score 23; DB 2; Length 13146;
 Best Local Similarity 74.4%; Pred. No. 3.4;
 Matches 23; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 6 taagaatctcaaacaccccaaacacaaagttctcaaacaa 44
 |||
 DB 4429 TAGGAATCTTAAACAGCCTAACCAATTAATCAACAGCCA 491

RESULT 4
 US-09-270-984A-3/c
 : Sequence 3, Application US/09270984A
 : Patent No. 6048965
 : GENERAL INFORMATION:
 : APPLICANT: Dietz, Harry C.
 : TITLE OF INVENTION: MAMMALIAN REGULATOR OF
 : TITLE OF INVENTION: NONSENSE MEDIATED RNA DEGRAD
 : NUMBER OF SEQUENCES: 29
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Fish & Richardson, P.C.
 : STREET: 4225 Executive Square, Suite 1400
 : CITY: La Jolla
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 92037
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: Windows95
 : SOFTWARE: FastSeq for Windows Version 2.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US, 09, 270, 984A

US-09-312-266-3

Query Match 43.8%; Score 21; DB 4; Length 2640;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 40; Conservative 0; Mismatches 15; Indels

[illegible]

```

RECEIVED 8
US 506 543-673 1/c
: Sequence 1, Application US/08460673
: Patent No. 5479544
: GENERAL INFORMATION:
: APPLICANT: Flour, Reinhard
: APPLICANT: Flour, Alain
: APPLICANT: Yeh, Patrice
: TITLE OF INVENTION: MODIFIED KLOVEROMYCES YEASTS, THEIR
: TITLE OF INVENTION: PREPARATION AND USE
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Rhone-Poulenc Rorer Inc.,
: STREET: 500 Arcoleta Rd. #643
: CITY: Collegeville
: STATE: PA
: COUNTRY: USA
: ZIP: 19002
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In. Ref. no. #1.05, Version #1.05
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: 08460673

```

Query Match 43.38; Score 20.8; DB 1; Length 1685;
Best Local Similarity 70.08; Pred. No. 15;
Matches 29; Consensitive 0; Mismatches 12; Indels 0

OY 1 aaagctatagaatctcgaacaccaccaaacgagctctcaag 40
||| ||| ||| | 1111111 ||| | |||
db 1070 aaagcttgagatttaccacaccctaaacacatcttagcag 1081

9
RESULT
US-08-484-044-10/c
? Sequence 19, Affiliated US/08/484/044
? Patent No. 5552242
? GENERAL INFORMATION:
? APPLICANT: Caskey, C. T.
? APPLICANT: Fu, Yinghui
? APPLICANT: Friedman, David L.
? APPLICANT: Rizzuti, Antonio
? APPLICANT: Penwick, Raymond G.
? TITLE OF INVENTION: Diagnostics of Myocardial Necrosis
? NUMBER OF SEQUENCES: 14
? CORRESPONDENT ADDRESS:
? ADDRESSEE: Fulbright & Jaworski, L.L.P.
? STREET: 1301 McKinney, Suite 5100
? CITY: Houston
? STATE: Texas
? COUNTRY: U.S.A.
? ZIP: 77010-4095
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC DOS/MS-DOS
? SOFTWARE: Patent to Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/484-044

| | | | | |
|-----------------------|-----------------|----------------|-----------|---------------|
| Query Match | 43.38; | Score 20.8; | DB 1; | Length 11613; |
| Best Local Similarity | 70.08; | Prod. No. 23; | | |
| Matches 29; | Conservative 0; | Mismatches 12; | Indels 0; | |

Qy 2 aaactaaagtctcagaacaccacaacagatctcaqa 41
||| ||||| | ||||||| | ||||| |||
Hb 4005 AACCCAGAGGTCACCCACACCACAGAAATAGCTCCACA 2966

RESULT 10
US 08 181 271A 36
SEQUENCE NO. APPLICATION NO. 0808181271A
PATENT NO. 5614395
GENERAL INFORMATION:
APPLICANT: Ryals, John A.
APPLICANT: Alexander, Danny C.
APPLICANT: Beck, James J.
APPLICANT: Duesing, John B.
APPLICANT: Friedrich, Leslie B.

117: qb_est48:*
 118: qb_est49:*
 119: qb_est50:*
 120: qb_est51:*
 121: qb_est52:*
 122: qb_est53:*
 123: qb_est54:*
 124: qb_est55:*
 125: qb_est56:*
 126: qb_est57:*
 127: qb_est58:*
 128: qb_est59:*
 129: qb_est60:*
 130: qb_est61:*
 131: qb_est62:*
 132: qb_est63:*
 133: qb_est64:*
 134: qb_est65:*
 135: qb_est66:*
 136: qb_est67:*
 137: qb_est68:*
 138: qb_est69:*
 139: qb_est70:*
 140: qb_est71:*
 141: qb_est72:*
 142: qb_est73:*
 143: qb_est74:*
 144: qb_est75:*
 145: qb_est76:*
 146: qb_est77:*
 147: qb_est78:*
 148: qb_est79:*
 149: qb_est80:*
 150: qb_est81:*
 151: qb_est82:*
 152: qb_est83:*
 153: qb_est84:*
 154: qb_est85:*
 155: qb_est86:*
 156: qb_est87:*
 157: qb_est88:*
 158: qb_est89:*
 159: qb_est90:*
 160: qb_est91:*
 161: qb_est92:*
 162: qb_est93:*
 163: qb_est94:*
 164: qb_est95:*
 165: qb_est96:*
 166: qb_est97:*
 167: qb_est98:*
 168: qb_est99:*
 169: qb_est100:*
 170: qb_est101:*
 171: qb_est102:*
 172: qb_est103:*
 173: qb_est104:*
 174: qb_est105:*
 175: qb_est106:*
 176: qb_est107:*
 177: qb_est108:*
 178: qb_est109:*
 179: qb_est110:*
 180: qb_est111:*
 181: qb_est112:*
 182: qb_est113:*
 183: qb_est114:*
 184: qb_est115:*
 185: qb_est116:*
 186: qb_est117:*
 187: qb_est118:*
 188: qb_est119:*
 189: qb_est120:*

190: em_qss_pln1:*
 191: em_qss_pln2:*
 192: em_qss_pro1:*
 193: em_qss_pro2:*
 194: em_qss_rod1:*
 195: em_qss_rod2:*
 196: em_qss_rod3:*
 197: em_qss_rod4:*
 198: em_qss_rod5:*
 199: em_qss_vrt1:*
 200: em_qss_vrt2:*
 201: em_qss_vrt3:*
 202: qb_qss1:*
 203: qb_qss2:*
 204: qb_qss3:*
 205: qb_qss4:*
 206: qb_qss5:*
 207: qb_qss6:*
 208: qb_qss7:*
 209: qb_qss8:*
 210: qb_qss9:*
 211: qb_qss10:*
 212: qb_qss11:*
 213: qb_qss12:*
 214: qb_qss13:*
 215: qb_qss14:*
 216: qb_qss15:*
 217: qb_qss16:*
 218: qb_qss17:*
 219: qb_qss18:*
 220: qb_qss19:*
 221: qb_qss20:*
 222: qb_qss21:*
 223: qb_qss22:*
 224: qb_qss23:*
 225: qb_qss24:*
 226: qb_qss25:*
 227: qb_qss26:*
 228: qb_qss27:*
 229: qb_qss28:*
 230: qb_qss29:*
 231: qb_qss30:*
 232: qb_qss31:*
 233: qb_qss32:*
 234: qb_qss33:*
 235: qb_qss34:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | % Match | Query Length | DB | ID | Description |
|--------|------|-------|---------|--------------|----------|----------|-------------|
| 1 | 48 | 100.0 | 559 | 165 | BE276194 | HE276194 | 601144462 |
| 2 | 48 | 100.0 | 586 | 166 | BE390259 | HE390259 | 601285131 |
| 3 | 48 | 100.0 | 689 | 165 | BE274123 | HE274123 | 601120646 |
| 4 | 35.4 | 73.8 | 739 | 166 | BE391242 | HE391242 | 601285820 |
| 5 | 30.6 | 63.8 | 800 | 166 | BE408892 | HE408892 | 601303769 |
| 6 | 28.2 | 58.8 | 674 | 166 | BE388562 | HE388562 | 601281855 |
| c 7 | 26 | 54.2 | 400 | 151 | HE662245 | HE662245 | maad003 |
| c 8 | 26 | 54.2 | 415 | 148 | HE452889 | HE452889 | maad003 |
| c 9 | 26 | 54.2 | 475 | 146 | HE321797 | HE321797 | 4264109.7 |
| c 10 | 26 | 54.2 | 534 | 137 | HE635044 | HE635044 | uv75c07.Y |
| c 11 | 26 | 54.2 | 542 | 137 | HE632718 | HE632718 | uv75c07.Y |
| c 12 | 26 | 54.2 | 579 | 3 | AA163664 | AA163664 | uv75c07.Y |
| c 13 | 26 | 54.2 | 686 | 173 | RG081278 | RG081278 | int19d11.r |
| c 14 | 25.8 | 52.9 | 526 | 225 | AZ637601 | AZ637601 | H3053008 |
| c 15 | 25.8 | 53.9 | 640 | 218 | AZ284687 | AZ284687 | 1M0497803 |
| c 16 | 25.8 | 53.9 | 686 | 236 | AZ284687 | AZ284687 | RFC1_23-1 |
| c 17 | 25.2 | 52.5 | 467 | 193 | A1890099 | A1890099 | 0131604.X |
| c 18 | 25.2 | 52.5 | 688 | 220 | AZ381858 | AZ381858 | 1M0497803 |

Copyright (c) 1993 2000 Compugen Ltd.

Signature Version 4.5

Database: us-09-341-829a-4_copy_623_679

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Query: us-09-341-829a-4_copy_623_679
Search time: 2512.006 seconds
(without attachments)
444,528 Million coll updates/sec

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 57 | 100.0 | 217 | 9 | AK042532 Sequence |
| 2 | 57 | 100.0 | 994 | 9 | AX024609 Sequence |
| 3 | 57 | 100.0 | 994 | 92 | HS012835 Homo sapi |
| 4 | 57 | 100.0 | 1002 | 9 | AK042535 Sequence |
| 5 | 57 | 100.0 | 1002 | 92 | HS0224040 Homo sapi |
| 6 | 57 | 100.0 | 1245 | 91 | HS0224093 Homo sapi |
| 7 | 57 | 100.0 | 112874 | 78 | AF277315 Homo sapi |
| 8 | 55.4 | 97.2 | 874 | 92 | HS0275978 Homo sapi |
| 9 | 55.4 | 97.2 | 2630 | 92 | HS0275977 Homo sapi |
| 10 | 55.4 | 97.2 | 112874 | 78 | AF277315 Homo sapi |
| 11 | 27.2 | 47.7 | 2755 | 94 | AF029107 Rat10g107 |


```

DEFINITION Homo sapiens chromosome 16 clone RP11-291F1, LOW-PASS SEQUENCE.
ACCESSION AC010547.
VERSION 1.0.
KEYWORDS HIG: HIGS_PHASE0.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 214911)
AUTHORS DeF Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 214911)
AUTHORS DeF Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DeF Joint
Genome Institute, 2000 Mitchell Drive, Walnut Creek, CA 94598, USA
on Jan 26, 2000 this sequence version replaced q1:5892401.
COMMENT -----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
*****
* NOTE: This record contains 422 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 365: contig of 365 bp in length
* 366 gap of unknown length
* 1136: contig of 771 bp in length
* 1147 gap of unknown length
* 1940: contig of 794 bp in length
* 2606: gap of unknown length
* 3285: contig of 679 bp in length
* 3286 gap of unknown length
* 3873: contig of 588 bp in length
* 4874 gap of unknown length
* 4171: contig of 298 bp in length
* 4172 gap of unknown length
* 4913 gap of unknown length
* 5649: contig of 727 bp in length
* 5640 gap of unknown length
* 5698: contig of 258 bp in length
* 6459: contig of 562 bp in length
* 6460 gap of unknown length
* 7428: contig of 969 bp in length
* 7429 gap of unknown length
* 8036: contig of 598 bp in length
* 8027 gap of unknown length
* 8690: contig of 664 bp in length
* 8691 gap of unknown length
* 9440: contig of 749 bp in length
* 10052: gap of unknown length
* 10053 contig of 613 bp in length
* 10494: gap of unknown length
* 11481: contig of 442 bp in length
* 10495 gap of unknown length
* 11482 contig of 887 bp in length
* 11640: gap of unknown length
* 11641 contig of 249 bp in length
* 366 gap of unknown length
*
* 12456: contig of 826 bp in length
* 12457 gap of unknown length
* 12775: contig of 319 bp in length
* 13406: gap of unknown length
* 13407 contig of 531 bp in length
* 13408 gap of unknown length
* 13488: contig of 182 bp in length
* 13489 gap of unknown length
* 13489 contig of 161 bp in length
* 13650: gap of unknown length
* 14427: contig of 778 bp in length
* 15752: gap of unknown length
* 15753 contig of 125 bp in length
* 15754 gap of unknown length
* 15754 contig of 234 bp in length
* 16216: gap of unknown length
* 16216 contig of 730 bp in length
* 17500: gap of unknown length
* 17500 contig of 784 bp in length
* 18474: gap of unknown length
* 18474 contig of 874 bp in length
* 19470: gap of unknown length
* 19470 contig of 996 bp in length
* 19471 gap of unknown length
* 19471 contig of 488 bp in length
* 20049: gap of unknown length
* 20049 contig of 191 bp in length
* 20844: gap of unknown length
* 20844 contig of 795 bp in length
* 21641: gap of unknown length
* 21641 contig of 797 bp in length
* 22125: gap of unknown length
* 22125 contig of 484 bp in length
* 22126 gap of unknown length
* 22126 contig of 177 bp in length
* 23005: gap of unknown length
* 23005 contig of 703 bp in length
* 23006 gap of unknown length
* 23709: contig of 704 bp in length
* 23710 gap of unknown length
* 24895: contig of 1186 bp in length
* 25012: gap of unknown length
* 25012 contig of 117 bp in length
* 25853: gap of unknown length
* 25853 contig of 841 bp in length
* 26726: gap of unknown length
* 26726 contig of 873 bp in length
* 27416: gap of unknown length
* 27416 contig of 690 bp in length
* 27417 gap of unknown length
* 27417 contig of 718 bp in length
* 28125: gap of unknown length
* 28125 contig of 121 bp in length
* 28256: gap of unknown length
* 28256 contig of 258 bp in length
* 28514: gap of unknown length
* 28514 contig of 714 bp in length
* 29228: gap of unknown length
* 29228 contig of 465 bp in length
* 30843: gap of unknown length
* 30843 contig of 1151 bp in length
* 31551: gap of unknown length
* 31551 contig of 708 bp in length
* 31552 gap of unknown length
* 32114: contig of 763 bp in length
* 32115 gap of unknown length
* 32115 contig of 123 bp in length
* 32438: gap of unknown length
* 32438 contig of 358 bp in length
* 32796: gap of unknown length
* 32796 contig of 177 bp in length
* 32972: gap of unknown length
* 32972 contig of 177 bp in length
* 33261: gap of unknown length
* 33261 contig of 289 bp in length

```


The present sequence encodes LAGE-1 tumour associated protein (TAP).
 The present invention also describes: (1) a method for treating a
 subject with a disorder characterised by expression of a LAGE-1 nucleic
 acid molecule or an expression product, comprising administering to the
 subject antitopoisomerase cytotoxic T cells to ameliorate the disorder, where
 the cytotoxic T cells are specific for complexes of an HLA molecule and
 a LAGE-1 TAP or an immunogenic fragment; (2) a method for treating a
 subject with a disorder characterised by expression of a LAGE-1 nucleic
 acid molecule or an expression product, comprising administering a
 LAGE-1 TAP or an immunogenic fragment to ameliorate the disorder; and
 (3) a method for selectively enriching a population of T cells with
 cytotoxic T cells specific for a LAGE-1 TAP comprising contacting an
 isolated population of T cells with an agent presenting a complex of a
 LAGE-1 TAP or an immunogenic fragment and a HLA presenting molecule to
 selectively enrich the isolated population of T cells with the cytotoxic
 T cells; the methods and products from the present invention can be used
 for the diagnosis and treatment of LAGE-1 associated disorders,
 particularly leucemias.

| | |
|--|--|
| query Match | 100.0%; Score: 57; DB: 19; Length: 1002; |
| Best Local Similarity | 100.0%; Pred. No.: 6, 50, 12; |
| Matches: 57; Conserved: 0; Mismatch: 0; Indels: 0; | |

[illegible]

RESULT 4
 $\frac{c^1}{c^1 + 4c^0}$
 1) Z^1/Z^0 2) S^1 and 4) c^1 , 1) N^1 , 5) c^1 4) c^1 .

$$\begin{array}{l} \text{XX} \\ \text{X}^{\vee} \\ Z_0/Z_{\text{aff}}; \\ \text{Aff} \\ \text{Aff}^{+} \end{array}$$

XX Hepatitis C virus RNA sequence SEQ ID No:4,
XX
KW Hepatitis C virus; RNA virus; replication; viral infection; ds,

Repatitis virus.
XX
IN W 06799 AL.

[illegible]

XX XX
XX XX

74 JUN 06 9178Z.

XX XX

VWV (WIND) SIXTY KK.

| | |
|----|---|
| XX | |
| Xi | Kobayashi M., Kobayashi K., Fujita K., Matsuzaki J., Ohno T. H. |
| XX | |
| XX | WPI : Zang 106.06 Zing |

Vectors expressing full length p6 of HPA strains, used to clarify mechanisms of RNA viral replication, infection, and development remedies and therapeutics.

Example 1: Page 46; Japanese.

transcription of both terminals of the RNA virus genome. Also described is a method for screening drugs for inhibiting the replication of RNA viruses by using the RNA virus infected host cell and infected host cells.

With hepatitis A viral infection, the vector is useful in clarifying the mechanism of RNA viral replication, onset of RNA viral infection,

CC and developed remedies and the updates for RNA viral infections, particularly of hepatitis C virus, the present sequence represents a DNA sequence from hepatitis C virus, from an example of the present invention.

XX Sequence 5,37 bp; 88 A; 141 C; 125 G; 104 T; 0 other;

XX

| | | | | | | | |
|-----------------------|-------|--------------|----|------------|----|--------|-----|
| Query Match | 42.9% | Score | 25 | DB | 21 | Length | 637 |
| Best Local Similarity | 64.9% | Prod. No. | 4 | | | | |
| Matches | 47 | Conservative | 0 | Mismatches | 20 | Indels | 0 |

[illegible]

1D
XX
Z5/403 standard; DNA; 644 bp.
XX
Ac
Z57403;
XX

DE Hepatitis C virus DNA sequence SEQ ID NO:6.

XX Hepatitis C virus; replication; viral infection; 45.
XX Hepatitis C virus.

1N W09967394-A1,
 2X
 3D 29-14E*, 1999,
 4Y

| | | |
|----|-------------|----------------|
| FF | 24 JUN 1999 | 95W02-1104380. |
| XX | | |
| FR | 24 JUN 1998 | 98JF-0177820. |
| XX | | |

| | |
|----|---|
| FA | (0.005) + 0.0061 (SE) WKU KK. |
| XX | |
| PT | Kolita M, Kolita K, Laita K, Matso, Kij J, Chhetti B; |
| YY | |

BR Vectors expressing full length gene of RNA viruses, useful in
XX production mechanisms of RNA viral replication, infection and
PT

developing remedies and therapeutics.
Example 1: Page 47; Japanese.

CC The present invention describes a vector comprising a cDNA encoding an
CC 688A virus gene, constructed to ensure the exact and homogeneous
CC transcription of both terminals of the RNA virus gene. Also described
CC is a method for generating a library of recombinant cDNA clones.

CC virus by using the RNA viral infection model animal, particularly one with hepatitis C viral infection. The vector is useful in clarifying the mechanism of RNA viral replication, onset of RNA viral infection, and development of the disease in the RNA virus infection model.

CC particularly of a hepatitis C virus. The present sequence represents
 as a RNA sequence of a hepatitis C virus from an example of the present
 CC invention.

Sequence 644 BP; 107 A; 163 C; 160 G; 214 T; 0 other;

| | | |
|-----------------------|-------|---|
| Best Local Similarity | 64.98 | Prod. Ro. 4.1 |
| Matches | 37 | Conservative 0; Mismatches 20; Indels 0; gaps |

Table 9: Exact and approximate values of the eigenvalues of the operator \mathcal{L}_α for $\alpha = 0.5$.

```

RESULT      6
Q349049
ID   Z49049 standard; DNA; 1773 BP.
XX
AC   Z49049.
XX
DI   31-MAR-2000 (first entry)
XX
DE   HCV RNA polymerase coding sequence.
XX
KW   RNA polymerase, HCV, infection, therapy; ss.
XX
OS   Hepatitis C virus.
XX
PN   W09967396-A1.
XX
FD   29-SEP-1996
XX
PF   24-JUN-1999; 99WO-JP03381.
XX
PR   24-JUN-1998; 98JP 0177817.
XX
PA   (ITRE-) INT REAGENTS CORP.
XX
PA   (CHOS ) CHUGAI SEIYAKU KY
XX
PA   (TOYOZ) TOYODA T.
XX
PI   Toyoda T, Kohara M, Kohara K, Higashi K, Tsuchiya M.
XX
WI   WPI: 2000-106297/00
XX
IN   p-PSDB; Y56536
XX
PI   Polynucleotides encoding a hepatitis C virus (HCV)-derived RNA
XX
PI   polymerase, useful for screening polymerase inhibitors to treat HCV -
XX
PS   Disclosure; Page 14 17, 25pp, Japanese.
XX
CC   This sequence encodes the hepatitis C virus (HCV) RNA polymerase protein
XX
CC   of the invention. The protein can be used to detect inhibitors of RNA
XX
CC   polymerase, for use in the treatment of hepatitis C viral infections
XX
DB   Sequence 1773 BP; 412 A; 515 C; 461 G; 385 T; 0 other

Query Match      43.9%; Score 25; DB 21; Length 1773;
Best Local Similarity 64.9%; Pred. No. 3.8;
Matches 37; Conservative 0; Mismatches 20; Indels 0; Gaps 0.

QY   1 ccacccgaaggaagcccaaggaagatgggtgcagaagatgatccctttaaagtatgttc 57
DB   1483 ccaccttgagagctgcagacatcgggcccaagatgtccgcgttaagctgtgtcc 1539

RESULT      7
Q33282
ID   Q33282 standard; cDNA; 9472 BP.
XX
AC   Q33282.
XX
DI   19-MAY-1993 (first entry)
XX
DE   Korean hepatitis C virus full cDNA sequence LNC1.
XX
KW   HCV-LNC1; diagnosis; vaccine; ds.
XX
OS   Korean hepatitis C virus.
XX
FH   Key
XX
FT   CDS
XX
FT   343..9375
XX
FT   /*tag= a
XX
FT   /feature= "HCV polyprotein"
XX

```

```

PN   EP521318-A.
XX
PD   07-JAN-1993.
XX
FT   10 JUN 1992, 92EP 0109753.
XX
PR   10-JUN-1991; 91KP-0009510.
XX
PR   06-AUG-1991; 91KP-0013601.
XX
PA   (LUCK-) LUCKY LTD.
XX
XX
PI   Cho JM, Choi DY, Kim CH, Kim ST, Lee YB, Lim KJ, Park YW;
PI   So HS, Yang JY;
XX
XX
WI   WPI: 1993-001843/01
XX
DR   p-PSDB; R30616.
XX
XX
PI   DNA and polypeptide(s) from a new type of hepatitis C virus (HCV)
PI   for diagnosing and vaccinating against HCV infections
XX
XX
PS   Claim 7; Fig 2; 119pp; English.
XX
XX
CC   This is the full cDNA sequence of Korean hepatitis C virus (HCV)
XX
CC   cDNA, HCV-LNC1. Portions of the cDNA may be used as probes or
XX
CC   primers to diagnose the presence of the virus in putative samples.
XX
CC   Polypeptides derived from the cDNA may be used in a specific and
XX
CC   accurate method for detecting HCV antibodies in the serum of
XX
CC   hepatitis C patients. Antibodies directed against these polypeptides
XX
CC   are useful for the purification of HCV antigens and for the
XX
CC   development of an improved diagnostic to detect HCV antigens in a
XX
CC   sample. The polypeptides may also be used in a vaccine for treatment
XX
CC   and prevention of HCV infection at a dosage of 5-200 ug/peptide.
XX
XX
SQ   Sequence 9472 BP; 1897 A; 2829 C; 2396 G; 2073 T; 0 other;

```

```

Query Match      43.9%; Score 25; DB 14; Length 9472;
Best Local Similarity 64.9%; Pred. No. 5.3;
Matches 37; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```

```

QY   1 ccacccgaaggaagcccaaggaagatgggtgcagaagatgatccctttaaagtatgttc 57
DB   5629 ccaccttgagagctgcagacatcgggcccaagatgtccgcgttaagctgtgtcc 9139

```

```

RESULT      8
PI2467/c
ID   PI2467 standard; cDNA; 703 BP.
XX
AC   PI2467;
XX
DI   13-MAR-2001 (first entry)
XX
DE   Aspergillus oryzae EST SEQ ID NO:4990.
XX
XX
KW   Multiple gene expression; filamentous fungal cell; EST;
KW   expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW   Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW   culture condition; environmental stress; spore morphogenesis;
KW   metabolic pathway engineering; catabolic pathway engineering; ss.
XX
XX
OS   Aspergillus oryzae.
XX
XX
PN   W0200056762-A2.
XX
PD   28-SEP-2000
XX
XX
PF   22-MAR-2000; 2000WO-US07781.
XX
PR   22-MAR-1999; 99US-0273623.
XX
XX
PA   (NOVO ) NOVO NORDISK BIOTECH INC.
PA   (NOVO ) NOVO NORDISK AS.

```


PF 17-SEP-1990; 90EP-0310149.
XX
PR 21-DEC-1989; 89US-0456142.
PR 15-SEP-1989; 89US-0408045.
XX
PA (CHIR) CHIRON CORP.
PA (OYAA/) OYA A.
XX
PI Cha T., Han J., Houghton M., Irvine BD., Kolberg JA;
PI Miyamura T., Saito I., Weiner AJ;
XX
DR WPI: 1999-488843/41.
DR P-PSDB; Y14974.
XX
PI New Hepatitis C Virus isolates, useful for diagnosis of hepatitis
PI infections and development of vaccines
XX
PS Example 2; Fig 5; 122pp; English.
XX
CC The invention provides two new isolates of hepatitis C virus (HCV), J1
CC and J7. These two isolates comprise nucleotide and amino acid sequences
CC that are distinct from the HCV isolate HCV-1. The nucleotide sequences
CC may be used to detect non-A, non-B HCV (NANBH) polynucleotides by
CC hybridisation for diagnosis of NANBH infections. They may also be used to
CC screen blood donors, donated blood and blood products for this infection.
CC The isolates may also be used to isolate other naturally occurring
CC variants of the virus. The polypeptides may be used as a vaccine for
CC administration to patients to protect against infection with NANBH. The
CC present sequence represents the J1 NS5 domain coding sequence.
XX
SQ Sequence 427 BP; 91 A; 129 C; 105 G; 99 T; 3 other;

Query Match 41.1%; Score 22.4; DB 20; Length 427;
Best Local Similarity 63.2%; Prod. No. 11;
Matches 36; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 ccaaccgaaggaagcccaaggaagatggatgcagaaatgtatgccttttaattatgttgc 57
DB 345 ccaaccttcaagatctggaacacatcgaggcagaagatgtcgcgcctaaagctactgtcc 391

Search completed: May 1, 2001, 04:00:18
Job time: 7123 sec

[illegible]

117: qb_est48: *
 118: qb_est49: *
 119: qb_est50: *
 120: qb_est51: *
 121: qb_est52: *
 122: qb_est53: *
 123: qb_est54: *
 124: qb_est55: *
 125: qb_est56: *
 126: qb_est57: *
 127: qb_est58: *
 128: qb_est59: *
 129: qb_est60: *
 130: qb_est61: *
 131: qb_est62: *
 132: qb_est63: *
 133: qb_est64: *
 134: qb_est65: *
 135: qb_est66: *
 136: qb_est67: *
 137: qb_est68: *
 138: qb_est69: *
 139: qb_est70: *
 140: qb_est71: *
 141: qb_est72: *
 142: qb_est73: *
 143: qb_est74: *
 144: qb_est75: *
 145: qb_est76: *
 146: qb_est77: *
 147: qb_est78: *
 148: qb_est79: *
 149: qb_est80: *
 150: qb_est81: *
 151: qb_est82: *
 152: qb_est83: *
 153: qb_est84: *
 154: qb_est85: *
 155: qb_est86: *
 156: qb_est87: *
 157: qb_est88: *
 158: qb_est89: *
 159: qb_est90: *
 160: qb_est91: *
 161: qb_est92: *
 162: qb_est93: *
 163: qb_est94: *
 164: qb_est95: *
 165: qb_est96: *
 166: qb_est97: *
 167: qb_est98: *
 168: qb_est99: *
 169: qb_est100: *
 170: qb_est101: *
 171: qb_est102: *
 172: qb_est103: *
 173: qb_est104: *
 174: qb_est105: *
 175: qb_est106: *
 176: qb_est107: *
 177: qb_est108: *
 178: qb_est109: *
 179: qb_est110: *
 180: qb_est111: *
 181: qb_est112: *
 182: qb_est113: *
 183: qb_est114: *
 184: qb_est115: *
 185: qb_est116: *
 186: qb_est117: *
 187: qb_est118: *
 188: qb_est119: *
 189: qb_est120: *

190: em_oss_pln1: *
 191: em_oss_pln2: *
 192: em_oss_pro: *
 193: em_oss_rod1: *
 194: em_oss_rod2: *
 195: em_oss_rod3: *
 196: em_oss_rod4: *
 197: em_oss_rod5: *
 198: em_oss_vrt1: *
 199: em_oss_vrt2: *
 200: em_oss_vrt3: *
 201: qb_oss1: *
 202: qb_oss2: *
 203: qb_oss3: *
 204: qb_oss4: *
 205: qb_oss5: *
 206: qb_oss6: *
 207: qb_oss7: *
 208: qb_oss8: *
 209: qb_oss9: *
 210: qb_oss10: *
 211: qb_oss11: *
 212: qb_oss12: *
 213: qb_oss13: *
 214: qb_oss14: *
 215: qb_oss15: *
 216: qb_oss16: *
 217: qb_oss17: *
 218: qb_oss18: *
 219: qb_oss19: *
 220: qb_oss20: *
 221: qb_oss21: *
 222: qb_oss22: *
 223: qb_oss23: *
 224: qb_oss24: *
 225: qb_oss25: *
 226: qb_oss26: *
 227: qb_oss27: *
 228: qb_oss28: *
 229: qb_oss29: *
 230: qb_oss30: *
 231: qb_oss31: *
 232: qb_oss32: *
 233: qb_oss33: *
 234: qb_oss34: *

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | Query | Match | Length | DB | ID | Description |
|--------|-----|-------|-------|-------|--------|----|----------|---------------------|
| | 1 | 47.6 | 83.5 | 589 | 165 | | BE274124 | BE274123 601120646 |
| | 2 | 37.4 | 65.6 | 800 | 166 | | BE408892 | BE408892 601304769 |
| | 3 | 32.8 | 57.5 | 729 | 166 | | BE491232 | BE491232 601385920 |
| | 4 | 27.7 | 47.1 | 1141 | 171 | | BC232629 | BC232629 602357718 |
| c | 5 | 26.2 | 46.0 | 923 | 15 | | AT068383 | AT068383 meta0002a |
| c | 6 | 25.6 | 44.9 | 440 | 14 | | AI094027 | AI094027 est11610.x |
| c | 7 | 25.6 | 44.9 | 588 | 221 | | AZ411525 | AZ411525 IM0:84007 |
| c | 8 | 25.6 | 44.9 | 641 | 1 | | AA047981 | AA047981 mj26e10.r |
| c | 9 | 25.6 | 44.9 | 964 | 230 | | CNS020B4 | AL214248 Tetraodon |
| c | 10 | 25.6 | 44.9 | 1061 | 232 | | CNS051A9 | AL442321 Tetraodon |
| c | 11 | 25.2 | 44.2 | 831 | 134 | | AW350275 | AW350275 G8210007B |
| c | 12 | 24.9 | 43.5 | 1073 | 221 | | CNS052A9 | AL418942 Tetraodon |
| c | 13 | 24.5 | 43.2 | 728 | 171 | | BE966127 | BE966127 602286381 |
| c | 14 | 24.4 | 42.8 | 148 | 118 | | AW670670 | AW670670 114815 MA |
| c | 15 | 24.4 | 42.8 | 354 | 227 | | H48579 | H48579 RPO111-2018 |
| c | 16 | 24.2 | 42.5 | 424 | 223 | | AZ558236 | AZ558236 R921-22.2 |
| c | 17 | 24.2 | 42.5 | 590 | 138 | | BE702434 | BE702434 M92:NN111 |
| c | 18 | 24.2 | 42.5 | 955 | 150 | | BE575748 | BE575748 602135474 |


```

RE391242      749 bp      mRNA      EST      21-JUL-2000
LOCUS      60129592391 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:1607708 5',
DEFINITION      mRNA sequence.
ACCESSION      RE391242
VERSION      RE391242.1 GI:19436597
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 749)
AUTHORS      G.H.-M. Hui, J. Wu, J. Li, J. Q.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Published (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Tel: (401) 496-1550
              Email: Robert_Strausberg@nih.gov
              Tissue Procurement: ATCC
              cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LHC259 row: q column: 05
              High quality sequence stop: 631.
FEATURES             Location/Qualifiers
     source          1..749
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1607708"
                     /clone_lib="NIH_MGC_44"
                     /issue_type="endometrium, adenocarcinoma cell line"
                     /lab_host="HPLC (phage-resistant)"
                     /note="organ: uterus; Vector: pOTB7; Site:1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRV/XhoI sites using the following 5'
                     adaptor: GCGACGAG(G). Library constructed by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      121 a 213 c 283 g 122 t
ORIGIN
Query Match      57.5%; Score 32.8; DB 166; Length 749;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 47; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

CY      4 cccatgagatccagagatataatcagagatgctctttaa 47
DB      567 cgcaggagagacgcggagatgctcttccgagctgtttaa 610

RE391242      1144 bp      mRNA      EST      13-FEB-2001
LOCUS      602357719F1 NIH_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4486205 5',
DEFINITION      mRNA sequence.
ACCESSION      RE391242.1 GI:12753444
VERSION      RE391242.1
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 1144)
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Published (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Tel: (401) 496-1550

```

```

Email: Robert_Strausberg@nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM10329 row: c column: 06
High quality sequence stop: 623.
FEATURES             Location/Qualifiers
     source          1..1144
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4486205"
                     /clone_lib="NCI_CGAP_Mam1"
                     /issue_type="tumor, biopsy sample"
                     /dev_stage="3 months, vitrin"
                     /lab_host="H10K"
                     /note="organ: mammary; Vector: pCMV-Sport6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally. Primer: oligo dT.
                     Library constructed by Life Technologies. Investigation
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      297 a 245 c 337 g 265 t
ORIGIN
Query Match      47.4%; Score 27; DB 174; Length 1144;
Best Local Similarity 70.6%; Pred. No. 3;
Matches 36; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

CY      1 cccacccagagagacccagagatgagatccaaagatgctctttaa 51
DB      820 caacccagagagacccagagatgagatccaaagatgctctttaa 870

RE391242      923 bp      mRNA      EST      09-DEC-1999
LOCUS      6006243841 Magnaporthe oryzae Appressorium Stage cDNA library
DEFINITION      Magnaporthe oryzae cDNA clone made0024384; 5', mRNA sequence.
ACCESSION      AL068383
VERSION      AL068383.1 GI:3391458
KEYWORDS      EST.
SOURCE      Magnaporthe oryzae.
             Magnaporthe oryzae
             Eukaryota; Fungi; Ascomycota; Fungiomycelina; Sordariomycetes;
             Sordariomycetes Incertae sedis; Magnaportheaceae; Magnaporthe.
REFERENCE      1 (bases 1 to 923)
AUTHORS      Choi, W., Fung, E., Sasinowski, M., Wang, R. and Dean, R.A.
TITLE      Expressed sequence characterization during appressorium formation
             in rice blast fungus, Magnaporthe oryzae
JOURNAL      Unpublished (1998)
COMMENT      Contact: Dean, R.A.
             Clemson University Genomics Institute
             100 Jordan Hall, Clemson University, Clemson, SC 29634, USA
             Tel: 864 656 5737
             Fax: 864 656 4293
             Email: rdean@clemson.edu
             Seq primer: 13 primer (AATTACCCCTCACTAAAGG)
             High quality sequence stop: 154.
FEATURES             Location/Qualifiers
     source          1..923
                     /organism="Magnaporthe oryzae"
                     /strain="70-15"
                     /db_xref="taxon:448105"
                     /clone="made0024384"
                     /clone_lib="Magnaporthe oryzae Appressorium Stage cDNA
                     library"
                     /dev_stage="germinated conidia on appressorium-inductive
                     surface"

```




FASTA File

FASTA File: us-09-341-829a-4-inv.res

FASTA File: us-09-341-829a-4-inv.res

FASTA File: us-09-341-829a-4-inv.res

FASTA File: us-09-341-829a-4-inv.res

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

FASTA File

NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/791,495
 FILING DATE:

CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Van Amsterdam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE/DOCKET NUMBER: 10461/7005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 217 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: internal
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens

Initial Score = 42 Optimized Score = 96 Significance = 0.13
 Residue Identity = 48% Matches = 109 Mismatches = 98
 Gaps = 16 Conservative Substitutions = 0

```

60      70      80      90      100      110      120
GTCTCTCTCAGGACAAACAAATCAGGCGCCGACAAAGAACTGGGCTCTGGACATTCCTCTAGGG
130      140      150      160      170      180      190
GAGGAGGATGACCTAG GAAGGAGGATGACCTAGG--GCTGATCCATATTAAGGAGCTGAGACCT
200      210      220      230      240      250      260
AAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
270      280      290      300      310      320      330
-TGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
340      350      360      370      380      390      400
AAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG

```

5. US-09-341-829a-4' (1-1002)
 US-08-791-495-11 Sequence 11, Application US/08791495
 Sequence 11, Application US/08791495

GENERAL INFORMATION:
 APPLICANT: Leth, Bernard
 APPLICANT: Lucas, Sophie
 APPLICANT: De Smet, Charles
 APPLICANT: Godelaine, Daniele
 APPLICANT: Boon-Falleur, Thierry
 TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/791,495
 FILING DATE:

CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Van Amsterdam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE/DOCKET NUMBER: 10461/0005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 19 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: YES
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens

Initial Score = 19 Optimized Score = 19 Significance = -0.45
 Residue Identity = 100% Matches = 19 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0

```

50      60      70      80      90      100      110      120
CATCTAACCGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
120      130      140      150      160
TTCCCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG

```

6. US 09 341-829a-4' (1-1002)
 US-08-791-495-3 Sequence 3, Application US/08791495
 Sequence 3, Application US/08791495

GENERAL INFORMATION:
 APPLICANT: Leth, Bernard
 APPLICANT: Lucas, Sophie
 APPLICANT: De Smet, Charles
 APPLICANT: Godelaine, Daniele
 APPLICANT: Boon-Falleur, Thierry
 TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston

STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/08791495
FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
TELEPHONE: 617-720-3500
TELEFAX: 617-720-3500
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
FEATURE: Linear
MOLECULE TYPE: cDNA
HYPOHETICAL: No
ANTI-SENSE: YES
ORIGINAL SOURCE:
ORGANISM: Homo sapiens

| | | | | | |
|------------------|-----|----------------------------|----|------------|-------|
| Initial Score | 19 | Optimized Score | 19 | Similarity | -0.4% |
| Residue Identity | 70% | Matches | 19 | Mismatches | 0 |
| Gaps | 0 | Conservative Substitutions | 0 | | |

Sequence 14, Application US/08791495
GENERAL INFORMATION:
APPLICANT: Leith, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Beau-Fallout, Thierry
TITLE OF INVENTION: IL-1 100% SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wall, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210

Sequence 14, Application US/08791495
GENERAL INFORMATION:
APPLICANT: Leith, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Beau-Fallout, Thierry
TITLE OF INVENTION: IL-1 100% SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wall, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210

Sequence 14, Application US/08791495

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/08791495
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
TELEPHONE: 617-720-3500
TELEFAX: 617-720-3500
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single

ALTERNATE/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
TELEPHONE: 617-720-3500
TELEFAX: 617-720-3500
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
FEATURE: Linear
MOLECULE TYPE: cDNA
HYPOHETICAL: No
ANTI-SENSE: No
ORIGINAL SOURCE:
ORGANISM: Homo sapiens

| | | | | | |
|------------------|-----|----------------------------|----|------------|-------|
| Initial Score | 14 | Optimized Score | 14 | Similarity | -0.9% |
| Residue Identity | 70% | Matches | 14 | Mismatches | 6 |
| Gaps | 0 | Conservative Substitutions | 0 | | |

Sequence 10, Application US/08791495
GENERAL INFORMATION:
APPLICANT: Leith, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Beau-Fallout, Thierry
TITLE OF INVENTION: IL-1 100% SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wall, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210

| | | | | | |
|------------------|-----|-----|-----|-----|-----|
| Initial Score | 440 | 340 | 350 | 460 | 470 |
| Residue Identity | 70% | 70% | 70% | 70% | 70% |
| Gaps | 0 | 0 | 0 | 0 | 0 |

Sequence 10, Application US/08791495
GENERAL INFORMATION:
APPLICANT: Leith, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Beau-Fallout, Thierry
TITLE OF INVENTION: IL-1 100% SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wall, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210

Sequence 10, Application US/08791495
GENERAL INFORMATION:
APPLICANT: Leith, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Beau-Fallout, Thierry
TITLE OF INVENTION: IL-1 100% SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wall, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210

Sequence 10, Application US/08791495

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/08791495
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
TELEPHONE: 617-720-3500
TELEFAX: 617-720-3500
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single

Sequence 10, Application US/08791495
GENERAL INFORMATION:
APPLICANT: Leith, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Beau-Fallout, Thierry
TITLE OF INVENTION: IL-1 100% SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wall, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210

Sequence 10, Application US/08791495
GENERAL INFORMATION:
APPLICANT: Leith, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Beau-Fallout, Thierry
TITLE OF INVENTION: IL-1 100% SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wall, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210

MOLECULE TYPE: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens

Initial Score - 9 Optimized Score - 10 Significance - -0.70
 Residue Identity - 55% Matches - 10 Mismatches - 8
 Gaps - 0 Conservative Substitutions - 0

90 100 110 120 130 140 150 X
 CACCAAAAGAAATGGCCATCTGCTGTGGTAAATATCAAGAGAGAGCAAGCACTAGAGAGAGAA
 | | | | |
 GCCATGCAGCGCGAAGGC X
 X 10

160 170 180 190 200
 CAGTCTGAGTTTAAATAGCTGTGCTGAGGAGAGCTGATGAGCAAAAA

9. US-09-341-829A-4' (1-1002)
 US-08-791-495-12 Sequence 12, Application US/08791495

Sequence 12, Application US/08791495

GENERAL INFORMATION:
 APPLICANT: Lucis, Bernard
 APPLICANT: Lucis, Sophie
 APPLICANT: De Smet, Charles
 APPLICANT: Godolaine, Daniele
 APPLICANT: Boon-Fallour, Thierry
 TITLE OF INVENTION: 12-1 TUMOR SPECIFIC GENES
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wall, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08791495
 FILING DATE:

CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Van Amsterdam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE/KEYWORD NUMBER: 16461/7005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 17 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens

Initial Score - 8 Optimized Score - 11 Significance - -0.73
 Residue Identity - 64% Matches - 11 Mismatches - 6
 Gaps - 0 Conservative Substitutions - 0

580 590 600 610 620 630 640 X

GGGCTTCGCGGAGAGAGTCTGGCATCCCGGCAAGCAATCTCTGGAGTAGTTCGTTCCATCGGAGAA
 | | | | | | | | | |
 CCGAGCGCTTCGCGAGTG X
 X 10

650 660 670 680 690
 CAGAAAAGCATCTGATGCTGCAACTGAAGCAGCGGCTGTCAGAGC

10. US-09-341-829A-4' (1-1002)
 US-08-791-495-2 Sequence 2, Application US/08791495

Sequence 2, Application US/08791495

GENERAL INFORMATION:
 APPLICANT: Lucis, Bernard
 APPLICANT: Lucis, Sophie
 APPLICANT: De Smet, Charles
 APPLICANT: Godolaine, Daniele
 APPLICANT: Boon-Fallour, Thierry
 TITLE OF INVENTION: 12-1 TUMOR SPECIFIC GENES
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wall, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08791495
 FILING DATE:

CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Van Amsterdam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE/KEYWORD NUMBER: 16461/7005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-2441
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 16 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens

Initial Score - 7 Optimized Score - 12 Significance - -0.75
 Residue Identity - 66% Matches - 12 Mismatches - 6
 Gaps - 0 Conservative Substitutions - 0

540 550 560 570 580 X 590 600
 CACATAAAATAGTAGTTGGAGGACACGGTGAAGTCTTCAGAACAGCTCTTCATAGGCGGAGAGTGAAGA
 | | | | | | | | | |
 AGATGGGTGCACAGCGTGT X
 X 10

610 620 630 640
 TCCGGGACAGCATTCCTGGGACACCGCTCCGCTTCATAGGCGGAGAG

11. US-09-341-829A-4' (1-1002)
 US-08-791-495-12 Sequence 12, Application US/08791495

FASTA File Information

FASTA File Pathwise Comparison of Sequences

Release 0.4

Results File: us-09-341-829a-4.res made by sdsaid on Tue 1 May 101 9:24:50 PDT.

Query sequence being compared: US 09-341-829A 4 (1-1002)

Number of sequences searched: 11

Number of sequences above cutoff: 11

Results of the initial comparison of US 09-341-829A 4 (1-1002) with:

File: US08791495.seq

100

N

U

M

E

R

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

E

PARAMETERS

Similarity matrix: Unitary: K: 1.00
Mismatch penalty: 1.00
Gap penalty: 0.34
Gap size penalty: 0.34
Gap size: 1
Randomization group: 0

SEARCH STATISTICS

Sequences: Mean: 298
Median: 19
Standard deviation: 426.45
Total elapsed: 00:00:00.00

Number of residues: 2958
Number of sequences searched: 11
Number of sequences above cutoff: 11

The scores below are sorted by initial score.
Similarity is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name: Description: Length: Score: Init. opt. Frame

1. US-08-791-495-4 Sequence 4, Application US 1002 1002 1002 2.45 0

The list of other best scores is:

| Sequence Name | Description | Length | Score | Init. opt. | Frame |
|----------------------|-----------------------------|--------|-------|------------|-------|
| 2. US-08-791-495-6 | Sequence 6, Application US | 755 | 518 | 574 | 0 |
| 3. US-08-791-495-8 | Sequence 8, Application US | 755 | 474 | 540 | 0 |
| 4. US-08-791-495-1 | Sequence 1, Application US | 217 | 217 | 217 | 0 |
| 5. US-08-791-495-12 | Sequence 12, Application US | 18 | 18 | 18 | 0 |
| 6. US-08-791-495-10 | Sequence 10, Application US | 18 | 18 | 18 | 0 |
| 7. US-08-791-495-2 | Sequence 2, Application US | 18 | 18 | 18 | 0 |
| 8. US-08-791-495-14 | Sequence 14, Application US | 20 | 9 | 11 | 0 |
| 9. US-08-791-495-13 | Sequence 13, Application US | 17 | 8 | 9 | 0 |
| 10. US-08-791-495-3 | Sequence 3, Application US | 19 | 8 | 13 | 0 |
| 11. US-08-791-495-11 | Sequence 11, Application US | 19 | 7 | 11 | 0 |

1. US-09-341-829A 4 (1-1002)
US-08-791-495-4 Sequence 4, Application US/08791495

Sequence 4, Application US/08791495

GENERAL INFORMATION:

APPLICANT: Leith, Bernard

APPLICANT: Lucas, Sophie

APPLICANT: De Smedt, Charles

APPLICANT: Godelaine, Danielle

APPLICANT: Boon-Pallour, Thierry

TITLE OF INVENTION: IL-1 1006R SPECIFIC GENES

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wolf, Greenfield & Sacks, P.C.

STREET: 600 Atlantic Avenue

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02210

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC DOS/MS DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 95/08791495

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Van Amsterdam, John R.

REGISTRATION NUMBER: 40,212

REFERENCE/DOCKET NUMBER: 10451/7005

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-720-3500

TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1002 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: 65..697

Initial Score: 1002
Residue Identity: 100%
Gaps: 0

1002
Matches
Conservative Substitutions

1002
Significance
2.45
0
0

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 09/08791495
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Van Amsterdam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE: P. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100
 TELEPHONE: 617-720-4441
 TELEFAX: 617-720-4500
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 18 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHEICAL: NO
 ANTI SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens

Initial Score: 18 Optimized Score: 18 Similarity: 0.58
 Residue Identity: 100% Matches: 18 Mismatches: 0
 Gaps: 0
 Conservative Substitutions: 0
 Sequence 1: Application US/08791495
 GENERAL INFORMATION:
 APPLICANT: Lech, Bernard
 APPLICANT: Lucas, Sophie
 APPLICANT: De Smet, Charles
 APPLICANT: Godelaine, Danielle
 APPLICANT: Boon-Fallour, Thierry
 TITLE OF INVENTION: EL-1 TUMOR SUPPRESSOR GENE
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Woll, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 09/08791495
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Van Amsterdam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE: P. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100
 TELEPHONE: 617-720-4441
 TELEFAX: 617-720-4500
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 18 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHEICAL: NO

6. US 09-341-829A-4 (1 1002)
 US 08-791-495-2 Sequence 10, Application US/08791495

Sequence 10, Application US/08791495
 GENERAL INFORMATION:
 APPLICANT: Lech, Bernard
 APPLICANT: Lucas, Sophie
 APPLICANT: De Smet, Charles
 APPLICANT: Godelaine, Danielle
 APPLICANT: Boon-Fallour, Thierry
 TITLE OF INVENTION: EL-1 TUMOR SUPPRESSOR GENE
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Woll, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 09/08791495
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Van Amsterdam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE: P. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100
 TELEPHONE: 617-720-4441
 TELEFAX: 617-720-4500
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 18 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHEICAL: NO

REFERENCE: P. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100
 TELEPHONE: 617-720-4441
 TELEFAX: 617-720-4500
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 18 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHEICAL: NO
 ANTI SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens

Initial Score: 18 Optimized Score: 18 Similarity: 0.58
 Residue Identity: 100% Matches: 18 Mismatches: 0
 Gaps: 0
 Conservative Substitutions: 0

Sequence 2, Application US/08791495
 GENERAL INFORMATION:
 APPLICANT: Lech, Bernard
 APPLICANT: Lucas, Sophie
 APPLICANT: De Smet, Charles
 APPLICANT: Godelaine, Danielle
 APPLICANT: Boon-Fallour, Thierry
 TITLE OF INVENTION: EL-1 TUMOR SUPPRESSOR GENE
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Woll, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 09/08791495
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Van Amsterdam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE: P. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100
 TELEPHONE: 617-720-4441
 TELEFAX: 617-720-4500
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 18 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHEICAL: NO

Sequence 2, Application US/08791495
 GENERAL INFORMATION:
 APPLICANT: Lech, Bernard
 APPLICANT: Lucas, Sophie
 APPLICANT: De Smet, Charles
 APPLICANT: Godelaine, Danielle
 APPLICANT: Boon-Fallour, Thierry
 TITLE OF INVENTION: EL-1 TUMOR SUPPRESSOR GENE
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Woll, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 09/08791495
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Van Amsterdam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE: P. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100
 TELEPHONE: 617-720-4441
 TELEFAX: 617-720-4500
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 18 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHEICAL: NO


```

PS  Column 4: Column 8; 8pp; English.
XX
CC  A novel cytoplasmic domain sequence (R93926) was identified in an
CC  integrin subunit beta-3 variant, termed beta-3' (see also R93929).
CC  This cytoplasmic domain or its subsequences (R93927-28) are used
CC  to produce anti-integrin antibodies useful in the analysis of
CC  cellular integrin content, in the determination of tumour origin,
CC  and in quantifying the complement of adhesion receptors on cells.
XX
SQ  Sequence 14 AA;

Query Match 77.8%; Score 28; DB 17; Length 14;
Best Local Similarity 84.4%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 0; Indels 6; Gaps 0;

QY  1 REGAGR 6
DB  142111
DB  2 rdaagr 7

RESULT 2
R93927
ID  R93927 standard; Peptide; 17 AA.
XX
AC  R93927;
XX
DI  20-MAY-1996 (first entry)
XX
DE  Integrin subunit beta-3' cytoplasmic domain (aa13-29).
XX
KW  Integrin subunit beta-3' cytoplasmic domain; antibody; tumour.
XX
OS  Homo sapiens.
XX
PN  J5543664-A.
XX
PD  12-MAR-1996.
XX
PF  25-MAY-1989; 890S-0357024.
XX
PR  25-MAY-1989; 890S-0357024.
XX
PR  09-NOV-1992; 920S-0974547.
XX
PR  10-MAY-1994; 940S-0240967.
XX
PA  (LJBL) LA JOLLA CANCER RES FOUND.
XX
PI  Roostahlt; EI;
XX
DI  WPI; 1996-159740/16.
XX
PT  Peptide comprising cytoplasmic domain of integrin subunit beta-3'
PT  used to produce anti-integrin antibodies, e.g. for cell integrin
PT  content analysis for the determination tumour origin
XX
PS  Claim 2; Column 8; 8pp; English.
XX
CC  A novel cytoplasmic domain sequence (R93926) was identified in an
CC  integrin subunit beta-3 variant, termed beta-3' (see also R93929).
CC  This cytoplasmic domain or its subsequences (R93927-28) are used
CC  to produce anti-integrin antibodies useful in the analysis of
CC  cellular integrin content, in the determination of tumour origin,
CC  and in quantifying the complement of adhesion receptors on cells.
XX
SQ  Sequence 17 AA;

Query Match 77.8%; Score 28; DB 17; Length 17;
Best Local Similarity 84.4%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  1 REGAGR 6
DB  142111
DB  2 rdaagr 7

RESULT 4
W23942
ID  W23942 standard; peptide; 26 AA.
XX
AC  W23942;
XX
DI  14-MAY-1998 (first entry)
XX
DE  Human cytoplasmic domain of the integrin beta3' subunit peptide.
XX
KW  human; integrin; beta3' subunit; cytoplasmic domain; antibody;
XX  immunoassay.
XX
OS  Homo sapiens.
XX

```

```

DB  6 rdaagr 11

RESULT 3
R93926
ID  R93926 standard; Peptide; 26 AA.
XX
AC  R93926;
XX
DI  20-MAY-1996 (first entry)
XX
DE  Integrin subunit beta-3' cytoplasmic domain (aa4-29).
XX
KW  Integrin subunit beta-3' cytoplasmic domain; antibody; tumour.
XX
OS  Homo sapiens.
XX
PN  J5543664-A.
XX
PD  12-MAR-1996.
XX
PF  25-MAY-1989; 890S-0357024.
XX
PR  25-MAY-1989; 890S-0357024.
XX
PR  09-NOV-1992; 920S-0974547.
XX
PR  10-MAY-1994; 940S-0240967.
XX
PA  (LJBL) LA JOLLA CANCER RES FOUND.
XX
PI  Roostahlt; EI;
XX
DI  WPI; 1996-159740/16.
XX
PT  Peptide comprising cytoplasmic domain of integrin subunit beta-3'
PT  used to produce anti-integrin antibodies, e.g. for cell integrin
PT  content analysis for the determination tumour origin
XX
PS  Claim 1; Column 7-8; 8pp; English.
XX
CC  A novel cytoplasmic domain sequence (R93926) was identified in an
CC  integrin subunit beta-3 variant, termed beta-3' (see also R93929).
CC  This cytoplasmic domain or its subsequences (R93927-28) are used
CC  to produce anti-integrin antibodies useful in the analysis of
CC  cellular integrin content, in the determination of tumour origin,
CC  and in quantifying the complement of adhesion receptors on cells.
XX
SQ  Sequence 26 AA;

Query Match 77.8%; Score 28; DB 17; Length 26;
Best Local Similarity 83.3%; Pred. No. 47;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  1 REGAGR 6
DB  15 rdaagr 20

RESULT 4
W23942
ID  W23942 standard; peptide; 26 AA.
XX
AC  W23942;
XX
DI  14-MAY-1998 (first entry)
XX
DE  Human cytoplasmic domain of the integrin beta3' subunit peptide.
XX
KW  human; integrin; beta3' subunit; cytoplasmic domain; antibody;
XX  immunoassay.
XX
OS  Homo sapiens.
XX

```

```

FN 0857-0664 A.
XX
XX 24 DEC 1997.
XX
XX 01 JUN 1999; 940S 0459246.
XX
XX 25 MAY 1999; 940S 045924.
XX 09 NOV 1999; 940S 0974547.
XX 10 MAY 1999; 940S 0240667.
XX 01 JUN 1999; 940S 0459246.
XX
XX (L142) LA P-LLA CANCER RES FOUND.
XX
XX P-051411 ELI.
XX
XX WPI: 1998 06243206.
XX
XX Integritin subunit beta 4' cytoplasmic domain peptide - useful for
XX producing antibodies for integrin immunoassays
XX
XX Claim 1; Volume 6; 7pp; English.
XX
XX The present sequence represents a peptide of the cytoplasmic domain
XX of integrin subunit beta 4'. The peptide can be used for producing
XX antibodies for use in immunoassays for integrins.
XX
XX Sequence 26 AA:

Query Match 77.8%; Score 28; DB 19; Length 26;
Best Local Similarity 84.6%; Pred. No. 47;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESACR 6
DB 18 rdqdr 24
L1111
18 rdqdr 24

RESULT 6
Y44474
ID Y44474 standard; peptide; 9 AA.
XX
XX Y44474;
XX
XX 27-MAR-2000 (first entry)
DI
DE Human antithrombin 111 variant 14.A (residues 485-499).
XX
XX Human; antithrombin 111; AT111 variant 14.A; elastase resistant;
XX IgG activated neutrophil resistant; anti-thrombin activity; heparin;
XX anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
XX thrombin activation-related pathological symptom; restenosis; thrombosis;
XX acute respiratory distress syndrome; thromboembolism; revascularization.
XX
XX Homo sapiens.
XX Synthetic.
XX
XX Key location/qualifiers
XX
XX FT Misc difference 5
XX
XX FT Misc difference 6
XX
XX FT Misc difference 6
XX
XX W09958098-A2.
XX
XX 18-NOV-1999.
XX
XX 12 MAY 1999; 99W0 0810549.
XX
XX 12 MAY 1998; 980S 0085197.
XX 05-MAY-1999; 940S-0085197.
XX
XX (3-7K/2) PAGE S.C.
XX (PICAF) PICARD V.
XX (ZENE) ZENEHROTH P
XX
XX Bock SC, Picard V, Zendeck-Roth P;
XX
XX WPI: 2000-116274/10
XX
XX New modified human antithrombin 111 compounds, used for treating e.g.
XX sepsis, trauma, acute respiratory distress syndrome, restenosis,
XX thrombosis, thromboembolism or stroke
XX
XX Claim 13; Page 57; 75pp; English.
XX
XX The present sequence is from an antithrombin 111 (AT111) variant, 14.A

```

```

FN 0857-0664 A.
XX
XX 24 DEC 1997.
XX
XX 01 JUN 1999; 940S 0459246.
XX
XX 25 MAY 1999; 940S 045924.
XX 09 NOV 1999; 940S 0974547.
XX 10 MAY 1999; 940S 0240667.
XX 01 JUN 1999; 940S 0459246.
XX
XX (L142) LA P-LLA CANCER RES FOUND.
XX
XX P-051411 ELI.
XX
XX WPI: 1998 06243206.
XX
XX Integritin subunit beta 4' cytoplasmic domain peptide - useful for
XX producing antibodies for integrin immunoassays
XX
XX Claim 1; Volume 6; 7pp; English.
XX
XX The present sequence represents a peptide of the cytoplasmic domain
XX of integrin subunit beta 4'. The peptide can be used for producing
XX antibodies for use in immunoassays for integrins.
XX
XX Sequence 26 AA:

Query Match 77.8%; Score 28; DB 19; Length 26;
Best Local Similarity 84.6%; Pred. No. 47;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RESACR 6
DB 18 rdqdr 24
L1111
18 rdqdr 24

RESULT 6
Y44474
ID Y44474 standard; peptide; 26 AA.
XX
XX Y44474.
XX
XX 20 MAY 1999. (first entry)
DI
DE Integritin subunit beta 4' cytoplasmic domain c-terminal sequence.
XX
XX Integritin subunit beta 4' cytoplasmic domain; antibody; tumour.
XX
XX Homo sapiens.
XX
XX 0857-0664 A.
XX
XX 12 MAR 1999.
XX
XX 25 MAY 1999; 940S 045924.
XX
XX 25 MAY 1999; 940S 045924.
XX 09 NOV 1999; 940S 0974547.
XX 10 MAY 1999; 940S 0240667.
XX
XX (L142) LA P-LLA CANCER RES FOUND.
XX
XX P-051411 ELI.
XX
XX WPI: 1998 159746/16.
XX N P80B; 116856.
XX
XX Peptide comprising cytoplasmic domain of integrin subunit beta-4'
XX used to produce anti-integrin antibodies, e.g. for cell integrin
XX content analysis for the determination tumour origin
XX

```

CC derived from human ATIII.N135A cDNA insert of the pBlueBac baculovirus
 CC expression construct and comprises residues 385-404. The variant has
 CC improved resistance to elastase and tpa-activated neutrophils while
 CC retains anti-thrombin and anti-factor Xa activities. It may be
 CC expressed as glycoforms with enhanced heparin affinity which target the
 CC blood vessel wall more efficiently than ATIIIs with normal heparin
 CC affinity. The modified ATIIIs can be used to treat thrombin activation-
 CC related pathological symptoms due to sepsis, trauma, acute
 CC respiratory distress syndrome, restenosis, thrombosis, thromboembolism
 CC and stroke. It can also be used to reduce the risk of reocclusion
 CC and restenosis in percutaneous transluminal coronary angioplasty,
 CC thrombolysis associated with surgery, ischaemia/reperfusion injury, and
 CC coagulation abnormalities in cancer or surgical patients.
 XX
 SQ Sequence: 9 AA:
 Query Match: 72.2%; Score 26; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 EGACR 6
 Db 5 exact 9
 RESULT 7
 Y44474
 ID Y44474 standard; peptide: 9 AA.
 AC Y44474:
 DT 27-MAR-2000 (first entry)
 DE Human antithrombin III variant 13.B (residues 385-404).
 KW Human; antithrombin III; ATIII variant 13.B; elastase-resistant;
 KW tpa-activated neutrophil resistant; anti-thrombin activity; heparin;
 KW anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
 KW thrombin activation-related pathological symptom; restenosis; thrombosis;
 KW acute respiratory distress syndrome; thromboembolism; reocclusion.
 OS Homo sapiens.
 CS Synthetic.
 EH Key: Location/Qualifiers
 FT Misc-difference 3 /note "ATIII.N135A Ala at 387 is substituted by Gly"
 FT Misc-difference 5 /note "ATIII.N135A Val at 389 is substituted by Gly"
 FT Misc-difference 6 /note "ATIII.N135A Ile at 390 is substituted by Gly"
 IN W0958098-A2.
 XX 18-NOV-1999.
 XX 12-MAY-1999; 99NO-US10549.
 XX 12-MAY-1998; 98US-0085197.
 XX 05-MAY-1999; 99US-0085197.
 XX (BCEP) ROCK S C.
 XX (PICA) PICARD V.
 XX (ZEHRO) ZENDEHROUH P.
 FT Rock SC; Picard V; Zendehtrouh P;
 IN WPI: 2000 116274/10
 XX New modified human antithrombin III compounds, used for treatment of:
 CC sepsis, trauma, acute respiratory distress syndrome, restenosis,
 CC thrombosis, thromboembolism or stroke.

XX
 PS Claim 13; Page 57; 75pp; English.
 CC The present sequence is from an antithrombin III (ATIII) variant, 13.B
 CC derived from human ATIII.N135A cDNA insert of the pBlueBac baculovirus
 CC expression construct and comprises residues 385-404. The variant has
 CC improved resistance to elastase and tpa-activated neutrophils while
 CC retains anti-thrombin and anti factor Xa activities. It may be
 CC expressed as glycoforms with enhanced heparin affinity which target the
 CC blood vessel wall more efficiently than ATIIIs with normal heparin
 CC affinity. The modified ATIIIs can be used to treat thrombin activation-
 CC related pathological symptoms due to sepsis, trauma, acute
 CC respiratory distress syndrome, restenosis, thrombosis, thromboembolism
 CC and stroke. It can also be used to reduce the risk of reocclusion
 CC and restenosis in percutaneous transluminal coronary angioplasty,
 CC thrombolysis associated with surgery, ischaemia/reperfusion injury, and
 CC coagulation abnormalities in cancer or surgical patients.
 XX
 SQ Sequence: 9 AA:
 Query Match: 72.2%; Score 26; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 EGACR 6
 Db 5 exact 9
 RESULT 8
 Y44475
 ID Y44475 standard; peptide: 9 AA.
 AC Y44475:
 DT 27-MAR-2000 (first entry)
 DE Human antithrombin III variant 13.C (residues 385-404).
 KW Human; antithrombin III; ATIII variant 13.C; elastase-resistant;
 KW tpa-activated neutrophil resistant; anti-thrombin activity; heparin;
 KW anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
 KW thrombin activation-related pathological symptom; restenosis; thrombosis;
 KW acute respiratory distress syndrome; thromboembolism; reocclusion.
 OS Homo sapiens.
 CS Synthetic.
 EH Key: Location/Qualifiers
 FT Misc-difference 3 /note "ATIII.N135A Ala at 387 is substituted by Gly"
 FT Misc-difference 4 /note "ATIII.N135A Val at 388 is substituted by Ile"
 FT Misc-difference 5 /note "ATIII.N135A Val at 389 is substituted by Gly"
 FT Misc-difference 6 /note "ATIII.N135A Ile at 390 is substituted by Gly"
 IN W0958098-A2.
 XX 18-NOV-1999.
 XX 12-MAY-1999; 99NO-US10549.
 XX 12-MAY-1998; 98US-0085197.
 XX 05-MAY-1999; 99US-0085197.
 XX (BCEP) ROCK S C.
 XX (PICA) PICARD V.
 XX (ZEHRO) ZENDEHROUH P.
 FT Rock SC; Picard V; Zendehtrouh P;

| | | | | | |
|----|--------|--|--|--|--|
| XX | | WP1: 2000 L16274.93 | | | |
| XX | | | | | |
| XX | | New modified human antithrombin III compounds, used for treating e.g. | | | |
| XX | P1 | sepsis, trauma, acute respiratory distress syndrome, restenosis, | | | |
| XX | P1 | thrombosis, thromboembolism or stroke. | | | |
| XX | P1 | | | | |
| XX | PS | claim 14; page 67; 75pp; English. | | | |
| XX | | | | | |
| XX | | The present sequence is from an antithrombin III (ATIII) variant, 13.C | | | |
| XX | CC | derived from human ATIII.R135A cDNA fragment of the phage baculovirus | | | |
| XX | CC | expression construct and comprises residues 385-393. The variant has | | | |
| XX | CC | improved resistance to elastase and tpa activated neutrophils while | | | |
| XX | CC | retains anti-thrombin and anti-factor Xa activities. It may be | | | |
| XX | CC | expressed as glycoproteins with enhanced heparin affinity which target the | | | |
| XX | CC | blood vessel wall more efficiently than ATIIIs with normal heparin | | | |
| XX | CC | affinity. The modified ATIIIs can be used to treat thrombin activation | | | |
| XX | CC | related pathological symptoms due to sepsis, trauma, acute | | | |
| XX | CC | respiratory distress syndrome, restenosis, thrombosis, thromboembolism | | | |
| XX | CC | and stroke. It can also be used to reduce the risk of reocclusion | | | |
| XX | CC | and restenosis in percutaneous transluminal coronary angioplasty, | | | |
| XX | CC | thrombotic associated with surgery, ischemia/reperfusion injury, and | | | |
| XX | CC | circulation abnormalities in cancer or surgical patients. | | | |
| XX | SC | Sequence: 9 AA; | | | |
| | | | | | |
| | | Query Match: 72.2%; Score: 26; EB 21; Length: 9; | | | |
| | | Best Local Similarity: 100.00%; Pred. No.: 3.2e+005; | | | |
| | | Matches: 5; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0; | | | |
| | | | | | |
| XX | DY | Z EAGLE # | | | |
| XX | DY | ELIIL | | | |
| XX | DY | S. EAGLE # | | | |
| | | | | | |
| XX | | RESULT: 9 | | | |
| XX | Y444.D | | | | |
| XX | DY | Y444.Y standard; peptide: 9 AA. | | | |
| XX | AC | Y444.Y; | | | |
| | | | | | |
| XX | | 27 MAP 2000 (first entry) | | | |
| XX | DY | Human antithrombin III variant 13.B (residues 385-393). | | | |
| XX | DE | | | | |
| XX | KW | Human: antithrombin III; ATIII variant 13.B; elastase-resistant; | | | |
| XX | KW | tpa activated neutrophil resistant; anti-thrombin activity; heparin; | | | |
| XX | KW | anti factor Xa activity; blood clotting disorder; sepsis, trauma, stroke; | | | |
| XX | KW | thrombin activation related pathological symptom; restenosis; thrombosis; | | | |
| XX | KW | acute respiratory distress syndrome; thromboembolism; reocclusion. | | | |
| XX | CS | Hemo sapiens. | | | |
| XX | CS | Synthetic. | | | |
| XX | XX | | | | |
| XX | FH | Key Location/Qualifiers | | | |
| XX | FT | Misc difference: 4 /note "ATIII.R135A Val at 388 is substituted by Leu" | | | |
| XX | F1 | Misc difference: 5 /note "ATIII.R135A Val at 389 is substituted by Glu" | | | |
| XX | F1 | Misc difference: 6 /note "ATIII.R135A Ile at 390 is substituted by Gly" | | | |
| XX | F1 | | | | |
| XX | PN | W0995R098 A2. | | | |
| XX | PD | 18 NOV 1999. | | | |
| XX | PF | 12 MAY 1999; 9909-USL0549. | | | |
| XX | TF | 12 MAY 1998; 9805-GBR05197. | | | |
| XX | PR | 05 MAY 1999; 9905-GBR05197. | | | |
| XX | XX | (HOWEKA) loc. S. C. | | | |

```

PR 18-DEC-1997; 97US-0070923.
PR 19-DEC-1997; 97US-0068169.
PR 19-DEC-1997; 97US-0068365.
PR 19-DEC-1997; 97US-0068367.
PR 19-DEC-1997; 97US-0068368.
PA (HUMAN) HUMAN GLOBIN SC1 INC.
XX
XX Carter KC, Duan RD, Feng P, Ferric AM, Florence C;
PI Florence K, Greene JM, Janat F, Kyaw H, Moore PA;
PI Ni J, Rosen CA, Ruben SM, Shi Y, Soppet DR, Wei Y;
PI Yu G;
XX WPI: 1999-418749/35
XX
XX New isolated human genes encoding secreted polypeptides
XX
XX Disclosure: Page 460; 537pp; English.
XX
XX X47916 to X48029 represent 110 isolated human secreted protein genes.
XX
XX Y46224 to Y46727 represent the secreted proteins encoded by the 110
XX human genes. The genes and their corresponding secreted polypeptides are
XX useful for preventing, treating or ameliorating medical conditions,
XX e.g. by protein or gene therapy. Also pathological conditions can be
XX diagnosed by determining the amount of the new polypeptides in a sample
XX or by determining the presence of mutations in the new genes. Specific
XX genes are described for each of the 110 genes, based on which tissues they
XX are most highly expressed in, and include developing products for the
XX diagnosis or treatment of cancer, tumors, developmental abnormalities
XX and fetal deficiencies, blood disorders, diseases of the immune system,
XX autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
XX disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
XX disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
XX disorders, digestive and other disorders, infections and AIDS. The
XX polypeptides are also useful for identifying their binding partners.
XX The sequences given in X47916 to X47919 and Y46224 are used in the
XX exemplification of the present invention.
XX
XX Sequence 26 AA:
SQ
Query Match 72.2%; Score 26; DB 20; Length 26;
Best Local Similarity 93.9%; Pred. No. 1,1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 RESAGR 6
DB 10 readmar 15
III II
XX
RESULT 11
Y52526
ID Y52526 standard; peptide; 25 AA.
XX
XX Y52526;
AC
XX
XX 22-FEB-2000 (first entry)
XX
XX House dust mite allergen 60 kD protein (mapd) N terminal fragment.
XX
XX Mite allergen protein; map: high molecular weight; HMW-map; allergy;
XX house dust mite; IgE; immunoglobulin E; allergen; mapd;
XX hypersensitivity reaction; therapy; treatment; diagnosis; human; feline;
XX canine; veterinary; antibody; vaccine; immunization.
XX
XX Dermatophagoides farinae.
XX
XX Key Location/Qualifiers
XX Misc-difference 1
XX /label- Xaa
XX /note- "Xaa - any amino acid"
XX
XX W09954349-A2.

```

```

XX 28-OCT-1999.
XX
XX 16-APR-1999; 99WO-US08524.
XX
XX 17-APR-1998; 98US-0062013.
XX 14 MAY 1998; 98US 0085295.
XX 02-SEP-1998; 98US-0098909.
XX
XX (HESK-) HESKA CORP.
XX
XX McCall CA, Hunter SW, Weber ER;
XX
XX WPI: 2000-052700/04.
XX
XX Novel high molecular weight Dermatophagoides nucleic acid polypeptides
XX used to modify an animal's hypersensitivity to mite allergens.
XX
XX Claim 3, Page 81, 154pp; English.
XX
XX This sequence represents an N-terminal fragment of the Dermatophagoides
XX fatiguae 60 kD mite allergen protein (mapd). The 60 kD allergen was
XX isolated from a 3.7 minute heparin-Sepharose gel filtration, with each
XX fraction being analysed for the presence of proteins that bound to IgE
XX present in allergic dog antisera. Mite allergenic proteins and
XX peptides, and nucleic acids encoding them, may be used in therapeutic
XX compositions to modify an animal's hypersensitivity reaction to mite
XX allergens. Animals that may be treated include mammals and birds,
XX especially felines, canines, equines, humans, other pets, and wild or
XX domestic animals. The proteins or fragments may also be used to
XX diagnose allergies via a skin test. The proteins and peptides can also
XX be used to raise antibodies, which have a variety of potential
XX uses. For example, they can be used as vaccines to passively immunize
XX animals against test mite hypersensitivity, as positive controls in a
XX test kit, and as tools to recover desired dust mite allergens from a
XX mixture of proteins.
XX
XX Sequence 25 AA:
SQ
Query Match 69.4%; Score 25; DB 21; Length 25;
Best Local Similarity 57.1%; Pred. No. 1,7e+02;
Matches 1; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 PEGAGRM 7
DB 17 rqaqkm 23
III III
XX
RESULT 12
R32448
ID R32448 standard; peptide; 10 AA.
XX
XX R32448;
AC
XX
XX 10-JUN-1993 (first entry)
XX
XX Human thrombospondin type I repeat derived peptide #21.
XX
XX thrombosis; coagulation; heparin binding inhibitor; type I repeat.
XX
XX Synthetic.
XX
XX D57801812-A.
XX
XX 15-DEC-1992.
XX
XX 06-DEC-1991, 91US 0801812.
XX
XX 06-DEC-1991; 91US-0801812.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX

```

P1 Roberts ID;
 XX WP1: 1993-06-74 (9/09).
 DR
 XX New sulphated glycoconjugate binding peptide(s) - from type 1
 P1 repeats of human thrombospondin, preventing interaction of the
 PT glycoconjugates with adhesion molecules, growth factors, etc.
 XX
 PS Disclosure: Page 13; 64pp; English.
 XX
 CC This peptide was obtained from the adhesive glycoprotein
 CC thrombospondin. It was tested for inhibitory activity against
 CC binding to heparin and laminin and was inactive. The preferred
 CC peptides of the invention (which do have inhibitory activity) all
 CC contain a subsequence WSXW (X: P, E, H, A, S) with a substantial
 CC lack of an electrical charge. They act as inhibitors of
 CC heparin or related sulphated glycoconjugate-binding to adhesion
 CC molecules, growth factors, etc.
 XX
 SQ Sequence: 10 AA;
 Query Match 66.7%; Score 24; DB 14; Length 10;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 4; Conservative 2; Mismatches 0; Gaps 0;
 QY 1 REGAGR 6
 DB 1 RKASR 6
 RESULT 14
 R02658
 ID R02658 standard; peptide: 10 AA.
 XX A*
 A* R02658;
 DE 18 AUG 2000 (first entry)
 DE Human thrombospondin derived peptide SEQ ID NO: 21.
 XX
 CC Human: thrombospondin; inhibitor; type 1 repeat unit; identification;
 CC extracellular matrix protein; heparin; proliferation; adhesion;
 CC metastasis; angiogenesis; neovascularisation; metastatic tumour;
 CC breast carcinoma; melanoma; cytotactin; heparin sulphate; RGF-2;
 CC fibroblast growth factor; 2; sulphate glycoconjugate; Kaposi's sarcoma,
 CC haemangioma; diabetic retinopathy.
 XX
 CS Homo Sapiens.
 CS Synthetic.
 XX
 XX US6051549 A
 DE 18 APR 2000
 XX
 XX 11 MAR 1998; 9405 0941119.
 XX
 XX 07 JUN 1995; 9405-0487568.
 XX 06 DEC 1994; 9405 0801812.
 XX 21 MAR 1994; 9405-0215085.
 XX
 CC (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 CC Chu N, Krut zsch HC, Roberts DE;
 XX WP1: 2000-44947/29.
 DR
 XX New polypeptides binding to heparin or related sulfated glycoconjugates
 CC are useful for inhibiting endothelial cell proliferation and tumor
 CC growth.
 XX
 XX Example 1: Column 18; 13pp; English.
 XX

CC The present invention describes a polypeptide (1) binding to heparin or
 CC related sulphated glycoconjugates with high affinity. R02638 to R02747
 CC represent peptides derived from human thrombospondin, which are used in
 CC the exemplification of the present invention. (1) can be used for
 CC inhibiting heparin or heparin sulphate interaction in a sample, without
 CC activating latent transforming growth factor (TGF)-beta, for inhibiting
 CC interaction of heparin or heparin sulphate with RGF-2 in a sample, and
 CC for inhibiting endothelial cell proliferation, and tumour growth, in
 CC patients. The peptides bind heparin or related sulphate glycoconjugates
 CC with high affinity. The peptides or their conjugates can be used in
 CC blocking or modifying the action on cellular processes of heparin (e.g.,
 CC proliferation, adhesion, motility, extracellular matrix remodelling),
 CC sulphatides, related sulphated glycoconjugates, fibronectin, and basic
 CC fibroblast growth factor, involving malignant cell lines and normal
 CC endothelial cells. The peptides can also be used in the treatment of
 CC metastatic tumours, breast carcinomas, melanomas, Kaposi's sarcomas,
 CC haemangiomas, diabetic retinopathies, and various pathological
 CC conditions dependent on neovascularisation.
 XX
 SQ Sequence: 10 AA;
 Query Match 66.7%; Score 24; DB 21; Length 10;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 4; Conservative 2; Mismatches 0; Gaps 0;
 QY 1 REGAGR 6
 DB 1 RKASR 6
 RESULT 14
 R79366
 ID R79366 standard; peptide: 19 AA
 XX AC
 XX R79366;
 DE 13-OCT-1995 (first entry)
 DE Thrombospondin heparin binding peptide Hept.
 XX
 CC Thrombospondin type 1 repeat sequence; transforming growth factor beta;
 CC wound healing; fibrosis; endothelial cell proliferation; heparin;
 CC Synthetic.
 XX
 XX W09505191-A.
 XX
 XX 23-FEB-1995.
 XX
 XX 12-AUG-1994; 94W0-0509194
 XX
 XX 13-AUG-1994; 94US-0106120.
 XX PR 04-MAY-1994; 94US-0248169.
 XX
 XX (DAHR-) DAB RES FOUND.
 XX
 XX Krut zsch HC, Murphy-Blirrich JE, Roberts DE, Schultz Cherry S;
 XX WP1: 1995-098579/13.
 XX
 CC Stimulating or inhibiting transforming growth factor beta by
 CC contacting with thrombo-spondin or an activating enzyme used
 CC to enhance wound healing or prevent fibrosis
 XX
 XX Examples, Page 28, 67pp, English.
 CC
 CC The sequence of a peptide lacking the thrombospondin heparin binding
 CC consensus motif: Trp-Ser-His-Tyr. The peptide corresponds to amino acid
 CC 17-35 of thrombospondin. The peptide was tested for the ability to
 CC stimulate the conversion of TGF beta from the latent to active form.
 CC Peptides (see R69766-79) which stimulate the conversion of latent TGF-beta
 CC to active TGF-beta can be used to enhance wound healing whilst inhibitory

CC peptides (R69780-90) can be used to prevent fibrosis or block TGF- β
 mediated endothelial cell proliferation.

SO Sequence 19 AA;

Query Match 66.7%; Score 24; DB 16; Length 19;

Best Local Similarity 66.7%; Posed No. 2002;

Mismatches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 REGAGR 6

II:III

Db 7 rkasqr 12

RESULT 15

R97641

II R97641 standard; peptide; 19 AA.

XX AC R97641;

XX 04-FEB-1997 (first entry)

XX DE Cell binding domain peptide RHD I of thrombospondin 1.

XX KW Thrombospondin 1; cell-binding domain; multidomain glycoprotein;

XX KW cell migration; cell attachment; cell differentiation; protein receptor;

XX KW integral membrane glycoprotein; cell adhesion; inflammation; arthritis;

XX KW angiogenesis; cancer metastasis.

XX OS Synthetic.

XX PN W9617059-A2.

XX PO 06-JUN-1996.

XX PF 21-NOV-1995; 95WO-0514553.

XX PR 30-NOV-1994; 94US-0347000.

XX PA (UNIW) UNIV WASHINGTON.

XX PI Frazier WA, Gao A;

XX PR 1986-277779/28.

XX PI Isolated thrombospondin 1 receptor protein - useful for studying

XX PI processes involved in inflammation, angiogenesis, arthritis and

XX PI cancer metastasis

XX PS Example: Fig 1; 43pp; English.

XX CC R97644-R97644 represent cell binding domain (CBD) peptides of

XX CC thrombospondin 1 (TSP1). Thrombospondins are a family of multidomain

XX CC glycoproteins that influence migration, attachment, and differentiation

XX CC of cell types. These CBB peptide sequences were used to identify the

XX CC protein of the invention. The protein of the invention is a 52 kD

XX CC protein receptor of the CBB peptides of TSP1. The receptor protein is an

XX CC integral membrane glycoprotein. The binding site of the protein receptor

XX CC is accessible on the intact cell surface, and the receptor mediates

XX CC adhesion of cells to the CBD of TSP1. Prior blockade of the protein

XX CC receptor by reaction of cells with the CBD peptides and a crosslinking

XX CC reagent inhibits cell adhesion. The receptor protein can be used to

XX CC study the migration, attachment, proliferation and differentiation of a

XX CC number of cell types. The receptor can also be used for assessing the

XX CC processes of inflammation, angiogenesis, arthritis and cancer metastasis.

XX SO Sequence 19 AA;

QY 1 REGAGR 6

II:III

Db 7 rkasqr 12

Search completed: May 1, 2001, 11:25:31

Job time: 101 sec

Query Match

Best Local Similarity 66.7%; Score 24; DB 17; Length 19;

Mismatches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 83.3%; Pred. No. 9;
Matches 5; Conservative 1; Mismatches 0;

QY 1 REGAGR 6 0, Gaps 0,
1:|||||
DB 6 RDGAGR 11

RESULT 3
5498694-9

Patent No. 5498694
APPLICANT: ROGSLAHTI, ERKKI I.
TITLE OF INVENTION: PEPTIDES OF THE CYTOPLASMIC DOMAIN OF
INTEGRIN

NUMBER OF SEQUENCES: 12
CURRENT APPLICATION DATA:
APPLICANT NUMBER: 93-06-24-567
FILING DATE: 10-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 973,547
FILING DATE: 09-NOV-1992
APPLICATION NUMBER: 357,824
FILING DATE: 25-MAY-1989

SEQ ID NO: 4
LENGTH: 26
5498694-9

Query Match 77.8%; Score 28; DB 6; Length 26;
Best Local Similarity 83.3%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 0; Indels 0, Gaps 0,

QY 1 REGAGR 6
1:|||||
DB 15 RDGAGR 20

RESULT 4
5498694-4

Patent No. 5498694
APPLICANT: ROGSLAHTI, ERKKI I.
TITLE OF INVENTION: PEPTIDES OF THE CYTOPLASMIC DOMAIN OF
INTEGRIN

NUMBER OF SEQUENCES: 12
CURRENT APPLICATION DATA:
APPLICANT NUMBER: 93-06-24-567
FILING DATE: 10-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 973,547
FILING DATE: 09-NOV-1992
APPLICATION NUMBER: 357,824
FILING DATE: 25-MAY-1989

SEQ ID NO: 4
LENGTH: 29
5498694-4

Query Match 77.8%; Score 28; DB 6; Length 29,
Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0, Gaps 0,

QY 1 REGAGR 6
1:|||||
DB 18 RDGAGR 23

RESULT 5
US-08-279-754-5

Sequence 5, Application US/08279754
Patent No. 5486470
GENERAL INFORMATION:
APPLICANT: DARKE, PAUL L.
APPLICANT: HALL, DAWN L.

APPLICANT: KUC, LAWRENCE C.
TITLE OF INVENTION: PURIFIED HERPES SIMPLEX VIRAL PROTEASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:

ADDRESSEE: JOANNE M. GIESSE
STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICANT NUMBER: US/28,279,754
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: GIESSE, JOANNE M.
REGISTRATION NUMBER: 32,838
REFERENCE/BOOK NUMBER: 19233

TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3046
TELEFAX: (908) 594-4720

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

HYPOTHETICAL: NO
FRAGMENT TYPE: internal
US-08-279-754-5

Query Match 69.4%; Score 25; DB 1; Length 14;
Best Local Similarity 71.4%; Pred. No. 29;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 REGAGR 7
1:|||||
DB 8 REGAGR 14

RESULT 6
US-08-687-820-4

Sequence 4, Application US/08687820
Patent No. 5972666
GENERAL INFORMATION:

APPLICANT: Hippenmeyer, Paul J
APPLICANT: Luckow, Verne A
APPLICANT: Rankin, Anne M
TITLE OF INVENTION: Assembly Deficient Herpesvirus Vaccine

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Joseph W. Bullock, G.D. Seale & Co.,
ADDRESSEE: Corporate Patent Dept.

STREET: P.O. Box 5110
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60680

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08,687,820
FILING DATE: 26-JUL-1996


```

1  CLASSIFICATION: 4.4
2  ALTERNATIVE INFORMATION:
3  NAME: Bolkow, Joseph W
4  REGISTRATION NUMBER: 47,193
5  REFERENCE NUMBER: C 2477
6  TELECOMMUNICATION INFORMATION:
7  TELEPHONE: 414.694.9094
8  TELEFAX: 414.694.9095
9  INFORMATION FOR SEQ ID NO: 4:
10 SEQUENCE CHARACTERISTICS:
11 LENGTH: 14 amino acids
12 TYPE: amino acid
13 SUPPLEMENTARY: single
14 topology: linear
15 MOLECULE TYPE:
16 DESCRIPTION: ZK 37 "SYNTHETIC"
17 US 09 341 829 A
18
19 Query Match: 66.4%, Score 25, ID 5, Length 14;
20 Best Local Similarity: 67.1%, Prod. No. 29;
21 Matches: 4, Conservative 4, Mismatches 0, Gaps 0;
22
23 QY 1 RESARCH 7
24 ID 1 KGSARV 7
25
26 RESULT 7
27 US 09 341 829 A
28 SEQUENCE 5, Application 10/TUS9609054
29 GENERAL INFORMATION:
30 APPLICANT: LARKE, PAUL G.
31 APPLICANT: HALL, DAWN L.
32 APPLICANT: KIL, LAWRENCE C.
33 TITLE OF INVENTION: PERFECTED HERPES SIMPLEX VIRAL PROLOGUE
34 NUMBER OF SEQUENCES: 5
35 CORRESPONDENCE ADDRESS:
36 ADDRESSEE: JAMES M. GLESSER
37 STREET: P.O. Box 2000, 126 E. LINCOLN AVE.
38 CITY: FAIRWAY
39 STATE: NJ
40 COUNTRY: USA
41 ZIP: 07035
42 COMPUTER READABLE FORM:
43 MEDIUM TYPE: floppy disk
44 COMPUTER: IBM PC compatible
45 OPERATING SYSTEM: PC-DOS/MS-DOS
46 SOFTWARE: PATENT IN PROCESS #1.0, Version #1.25
47 CURRENT APPLICATION DATA:
48 APPLICATION NUMBER: PAT/9505/00052
49 FILING DATE:
50 CLASSIFICATION:
51 ALTERNATIVE INFORMATION:
52 NAME: GLESSER, JEROME M.
53 REGISTRATION NUMBER: 42,849
54 REFERENCE NUMBER: C 19244
55 TELECOMMUNICATION INFORMATION:
56 TELEPHONE: (908) 594.4046
57 TELEFAX: (908) 594.4720
58 INFORMATION FOR SEQ ID NO: 5:
59 SEQUENCE CHARACTERISTICS:
60 LENGTH: 14 amino acids
61 TYPE: amino acid
62 SUPPLEMENTARY: single
63 topology: linear
64 MOLECULE TYPE: peptide
65 HYDROPHOBIC: No
66 INFORMATION:
67 US 09 341 829 A

```

Query Match: 66.4%, Score 25; ID 5; Length 14;

```

1  Best Local Similarity: 71.4%, Prod. No. 29;
2  Matches: 5, Conservative 1, Mismatches 0, Gaps 0;
3
4  QY 1 RESARCH 7
5  ID 8 RESARCH 14
6
7  RESULT 8
8  US-07-801-812A-21
9  SEQUENCE 21, Application US/07801812A
10 Patent No. 5357041
11 GENERAL INFORMATION:
12 APPLICANT: David D. Roberts et al
13 TITLE OF INVENTION: REPARTIN AND ANALOGUE BINDING
14 TITLE OF INVENTION: REPARTIN FROM THE TYPE 1 REPEATS OF HUMAN
15 TITLE OF INVENTION: THROUGHOUT IN FREEDOM MELANOMA CELL ADHESION
16 NUMBER OF SEQUENCES: 27
17 CORRESPONDENCE ADDRESS:
18 ADDRESSEE: Lowe, Pricco, Leblanc & Becker
19 STREET: Suite 300, 99 Canal Center Plaza
20 CITY: Alexandria
21 STATE: Virginia
22 COUNTRY: USA
23 ZIP: 22314
24 COMPUTER READABLE FORM:
25 MEDIUM TYPE: floppy disk
26 COMPUTER: IBM PC compatible
27 OPERATING SYSTEM: PC-DOS/MS-DOS
28 SOFTWARE: DOS TEXT FILE
29 CURRENT APPLICATION DATA:
30 APPLICATION NUMBER: US/07/801,812A
31 FILING DATE: 19911206
32 CLASSIFICATION: 530
33 ALTERNATIVE INFORMATION:
34 NAME: J.G. Mullins
35 REGISTRATION NUMBER: 44,074
36 REFERENCE NUMBER: 217-111
37 TELECOMMUNICATION INFORMATION:
38 TELEPHONE: 703.684.1111
39 INFORMATION FOR SEQ ID NO: 21:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 10 amino acids
42 TYPE: AMINO ACID
43 TOPOLOGY: linear
44 MOLECULE TYPE: peptide
45 US-07-801-812A-21
46
47 Query Match: 66.7%, Score 24; ID 1; Length 10;
48 Best Local Similarity: 66.7%, Prod. No. 35;
49 Matches: 4, Conservative 2, Mismatches 0, Gaps 0;
50
51 QY 1 RESARCH 6
52 ID 1 KGSARV 6
53
54 RESULT 9
55 US 08 487-568-21
56 SEQUENCE 21, Application US/08487568
57 Patent No. 5770564
58 GENERAL INFORMATION:
59 APPLICANT: Roberts, David D.
60 APPLICANT: Brown, Philip J.
61 APPLICANT: Bryant, Joseph L.
62 APPLICANT: Imman, John K.
63 APPLICANT: Kruttsch, Henry C.
64 APPLICANT: Guo, Nenghua
65 TITLE OF INVENTION: Repartin and Soluble Binding Peptides
66 TITLE OF INVENTION: from the Type 1 repeats of Human Transmembrane and
67 TITLE OF INVENTION: Computations thereof
68 NUMBER OF SEQUENCES: 114

```

1 CORRESPONDENCE ADDRESS:
 2 ADDRESSEE: Townsend and Townsend and Crew
 3 STREET: One Market Plaza, Sausalito Street Tower
 4 CITY: San Francisco
 5 STATE: California
 6 COUNTRY: USA
 7 ZIP: 94105-1492
 8
 9 COMPUTER READABLE FORM:
 10 MEDIUM TYPE: Floppy disk
 11 COMPUTER: IBM PC compatible
 12 OPERATING SYSTEM: PC-DOS/MS-DOS
 13 SOFTWARE: Patent In Release #1.0, Version #1.10
 14
 15 CURRENT APPLICATION DATA:
 16 APPLICATION NUMBER: US 08/215,085
 17 FILING DATE: 07-JUN-1995
 18 CLASSIFICATION: 514
 19
 20 PRIOR APPLICATION DATA:
 21 APPLICATION NUMBER: US 08/215,085
 22 FILING DATE: 21-MAR-1994
 23
 24 PRIOR APPLICATION DATA:
 25 APPLICATION NUMBER: US 07,991,872
 26 FILING DATE: 06-DEC-1991
 27 ATTORNEY/AGENT INFORMATION:
 28 NAME: Dow, Karel B.
 29 REGISTRATION NUMBER: 29,684
 30 REFERENCE/DOCKET NUMBER: 015290-023310
 31 TELEPHONE: (415) 326-2400
 32 TELEFAX: (415) 326-2422
 33 INFORMATION FOR SEQ ID NO: 21:
 34 SEQUENCE CHARACTERISTICS:
 35 LENGTH: 10 amino acids
 36 TYPE: amino acid
 37 STRANDEDNESS:
 38 TOPOLOGY: Linear
 39 MOLECULE TYPE: peptide
 40 US-08-487-569-21

Query Match: 66.7% Score 24; DB 1; Length 10;

Best local Similarity 66.7% Pred. No. 34;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 REGAGR 6
1 1111

DB 1 REGAGR 6

RESULT 10
 1 US-08-447-000-10
 2 Sequence 10; Application US/08447000
 3 Patent No. 5622665
 4 GENERAL INFORMATION:
 5 APPLICANT: Frazer, William A.
 6 APPLICANT: 54, A-342
 7 TITLE OF INVENTION: Receptor for Cell-binding Domain of
 8 TITLE OF INVENTION: Thrombospondins
 9 NUMBER OF SEQUENCES: 13
 10 CORRESPONDENCE ADDRESS:
 11 ADDRESSEE: Scott J. Meyer, Monsanto/Scarle, ASG
 12 STREET: 800 N. Lindbergh Blvd.
 13 CITY: St. Louis
 14 STATE: Missouri
 15 COUNTRY: USA
 16 ZIP: 63167
 17
 18 COMPUTER READABLE FORM:
 19 MEDIUM TYPE: Floppy disk
 20 COMPUTER: IBM PC compatible
 21 OPERATING SYSTEM: PC-DOS/MS-DOS
 22 SOFTWARE: Patent In Release #1.0, Version #1.25
 23
 24 CURRENT APPLICATION DATA:
 25 APPLICATION NUMBER: US 08/447,000
 26 FILING DATE:

1 CLASSIFICATION: 540
 2 PRIOR APPLICATION DATA:
 3 APPLICATION NUMBER: US 08/029,438
 4 FILING DATE: 05-MAR-1993
 5 ATTORNEY/AGENT INFORMATION:
 6 NAME: Meyer, Scott J.
 7 REGISTRATION NUMBER: 25,275
 8 REFERENCE/DOCKET NUMBER: WU-2848
 9 TELEPHONE: (314) 694-4117
 10 TELEFAX: (314) 694-5435
 11 INFORMATION FOR SEQ ID NO: 10:
 12 SEQUENCE CHARACTERISTICS:
 13 LENGTH: 19 amino acids
 14 TYPE: amino acid
 15 TOPOLOGY: Linear
 16 MOLECULE TYPE: peptide
 17 US-08-447-000-10

Query Match: 66.7% Score 24; DB 1; Length 19;
Best local Similarity 66.7% Pred. No. 64;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 REGAGR 6

DB 7 REGAGR 12

RESULT 11
 1 US-08-253-171-1
 2 Sequence 1; Application US/08253171
 3 Patent No. 5472863
 4 GENERAL INFORMATION:
 5 APPLICANT: MARGIA, Kazuhiko
 6 APPLICANT: KUBOTA, Michio
 7 APPLICANT: SUGIMOTO, Toshiyuki
 8 APPLICANT: MIYAKE, Toshio
 9 TITLE OF INVENTION: TROPALOSE-RELEASING ENZYME, AND ITS
 10 TITLE OF INVENTION: PREPARATION AND USES
 11 NUMBER OF SEQUENCES: 11
 12 CORRESPONDENCE ADDRESS:
 13 ADDRESSEE: BROWNY AND NEIMARK
 14 STREET: 419 Seventh Street, N.W., Suite 300
 15 CITY: Washington
 16 STATE: D.C.
 17 COUNTRY: USA
 18 ZIP: 20004
 19
 20 COMPUTER READABLE FORM:
 21 MEDIUM TYPE: Floppy disk
 22 COMPUTER: IBM PC compatible
 23 OPERATING SYSTEM: PC-DOS/MS-DOS
 24 SOFTWARE: Patent In Release #1.0, Version #1.25
 25
 26 CURRENT APPLICATION DATA:
 27 APPLICATION NUMBER: US/08/253,171
 28 FILING DATE: 02-JUN-1994
 29 CLASSIFICATION: 435
 30
 31 PRIOR APPLICATION DATA:
 32 APPLICATION NUMBER: JP 156338/1994
 33 FILING DATE: 03-JUN-1993
 34 PRIOR APPLICATION DATA:
 35 APPLICATION NUMBER: JP 340343/1994
 36 FILING DATE: 09-DEC-1993
 37
 38 PRIOR APPLICATION DATA:
 39 APPLICATION NUMBER: JP 79291/1994
 40 FILING DATE: 28-MAR-1994
 41 ATTORNEY/AGENT INFORMATION:
 42 NAME: NEIMARK, Sheridan
 43 REGISTRATION NUMBER: 20,520
 44 REFERENCE/DOCKET NUMBER: MARGIA 2
 45 TELECOMMUNICATION INFORMATION:
 46 TELEPHONE: 202-629-5197
 47 TELEFAX: 202-747-4528

TELEX: 248633
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US 09 341 829 1

Query Match 63.9% Score 23 DB 1 Length 10
 Best Local Similarity 80.0% Pred. No. 52
 Matches 4: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

QY 2 EAGR 6
 : : : :
 DB 5 QGAGR 9

RESULT 12
 US 09 443 352 1
 : Sequence 1, Application US/0844352
 : Patent No. 5591612

GENERAL INFORMATION:
 APPLICANT: MARUTA, Kazuhiko
 APPLICANT: KUROTA, Michio
 APPLICANT: SUGIMOTO, Toshiyuki
 APPLICANT: MIYAKE, Toshio
 TITLE OF INVENTION: TETRAHEDRE-RELEASING ENZYME, AND ITS
 TITLE OF INVENTION: PREPARATION AND USES
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEWMARK
 STREET: 419 Seventh Street, N.W., Suite 400
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20004

COMPUTER-RELEASABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC DOS/MS DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 09/0944352
 FILING DATE: 02 JUN 1994

CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/0944352
 FILING DATE: 02 JUN 1994

APPLICATION NUMBER: JP 79291/1994
 FILING DATE: 28 MAR 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: NEWMARK, Sheridan
 REGISTRATION NUMBER: 20,620
 REFERENCE/WORK NUMBER: MARUTA 2

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202 628-5197
 TELEFAX: 202 737-3528
 TELEX: 248633

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

MOLECULE TYPE: peptide
 US 09 443 352 1

Query Match 63.9% Score 23 DB 1 Length 10
 Best Local Similarity 80.0% Pred. No. 52
 Matches 4: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

QY 2 EAGR 6
 : : : :
 DB 5 QGAGR 9

RESULT 13
 US 09 443 352 1
 : Sequence 1, Application US/0844352
 : Patent No. 5591612

GENERAL INFORMATION:
 APPLICANT: MARUTA, Kazuhiko
 APPLICANT: KUROTA, Michio
 APPLICANT: SUGIMOTO, Toshiyuki
 APPLICANT: MIYAKE, Toshio
 TITLE OF INVENTION: TETRAHEDRE-RELEASING ENZYME, AND ITS
 TITLE OF INVENTION: PREPARATION AND USES
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWDY AND NEWMARK
 STREET: 419 Seventh Street, N.W., Suite 400
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20004

COMPUTER-RELEASABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC DOS/MS DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 09/0944352
 FILING DATE: 17 MAY 1995

CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/0944352
 FILING DATE: 02 JUN 1994

APPLICATION NUMBER: JP 79291/1994
 FILING DATE: 28 MAR 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: NEWMARK, Sheridan
 REGISTRATION NUMBER: 20,620
 REFERENCE/WORK NUMBER: MARUTA 2

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202 628-5197
 TELEFAX: 202 737-3528
 TELEX: 248633

INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US 09 443 352 1

Query Match 63.9% Score 23 DB 1 Length 10
 Best Local Similarity 80.0% Pred. No. 52
 Matches 4: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

QY 2 REAGR 6
 1111
 1 REAGR 9

RESULT 14
 5182464 10
 Patent No. 5182464
 TITLE OF INVENTION: POLYPEPTIDE ANALOGS OF APOBILIPROTEIN E
 NUMBER OF SEQUENCES: 14
 CURRENT APPLICATION DATA:
 APPLICANT: J. H. HEMER, US/08/485,158
 FILING DATE: 26-FEB-1990
 SEQ ID NO: 10
 LENGTH: 10
 5182464 10

Query Match 63.9%; Score 23; DB 6; Length 10;
 Best Local Similarity 80.1%; Pred. No. 52;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REAGR 6
 1111
 1 REAGR 6

RESULT 15
 US-08-471-7806-85
 Sequence 85, Application US/084717806
 Patent No. 5709608
 GENERAL INFORMATION:
 APPLICANT: Casterman, Cecile
 APPLICANT: Hamers, Raymond
 TITLE OF INVENTION: Immunoglobulins Devoid of Light Chains
 NUMBER OF SEQUENCES: 140
 CORRESPONDENCE ADDRESS:
 ADDRESS: Fimond, Henderson, Farabow, Garrett & Dunnet
 STREET: 1400 I Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/471,780C
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/106,944
 FILING DATE: 17-AUG-1993
 APPLICATION NUMBER: FR 92462426.0
 FILING DATE: 21-AUG 1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 93401410.3
 FILING DATE: 21-MAY-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Potter, Jane E.R.
 REGISTRATION NUMBER: 34,382
 REFERENCE/WORK NUMBER: 04958.0008-00000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-408-4000
 TELEFAX: 202-408-4400
 INFORMATION FOR SEQ ID NO: 85:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 14 amino acids
 TYPE: amino acid

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-471-7806-85

Query Match 63.9%; Score 23; DB 1; Length 14;
 Best Local Similarity 66.7%; Pred. No. 73;
 Matches 4; Conservative 1; Mismatches 0; Gaps 0;

QY 1 REAGR 6
 1111
 3 REAGR 8

Search completed: May 1, 2001, 11:27:42
 Job time: 152 sec


```

RESULT 4
SI:614
ribosomal protein l16 - Bacillus stearothermophilus (fragment)
C:Species: Bacillus stearothermophilus
C:Date: 30-Jun-1991 #sequence_revision 40-Jun-1991 #text_change 13-Aug-1999
C:Accession: S10614
R:Kromer, W.J.; Hatakeyama, T.; Kimura, M.
Biol. Chem. Hoppe-Seyler 371, 641-636, 1990
A:Title: Nucleotide sequences of Bacillus stearothermophilus ribosomal protein genes: pa
A:Reference number: S10610; M010:91025633
A:Accession: S10614
A>Status: not computed with conceptual translation
A:Molecule type: rRNA
A:Residues: 126 - 880.
A:Superfamily: eukaryotic ribosomal protein L16
C:Keywords: protein biosynthesis; ribosome

Query Match 63.9% Score 24; DB 2; Length 26;
Best Local Similarity 71.4%; Pred. No. 1,5e-02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 REWARD 7
DB 11 BHRCHM 17

RESULT 4
SI:2176
glutathione transferase (E.C. 2.5.1.18) alpha-Vx - rat (fragment)
C:Species: glutathione S-transferase Vx, glutathione transferase Yiotus
C:Date: 22-Nov-1994 #sequence_revision 30-Jan-1998 #text_change 30-Jan-1998
C:Accession: S21278
R:Kikuchi, T.; Ischiya, T.; Shikata, Y.; Sasaki, F.; Tadaya, O.; Horie, T.; Satoh, T.
Biochem. J. 283, 407-411, 1992
A:Title: Developmental aspects of a unique glutathione S-transferase subunit Vx in the
Yiotus.
A:Reference number: S21278, M010:92231842
A:Accession: S21278
A:Molecule type: protein
A:Residues: 1-28 - cDNA
C:Superfamily: glutathione transferase
C:Keywords: dimer; liver; transferase

Query Match 63.9% Score 26; DB 2; Length 28;
Best Local Similarity 66.7%; Pred. No. 1,5e-02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 REAGRM 7
DB 10 BHRCHM 15

RESULT 5
SI:9480
major immunophilin hsp56 - chicken (fragment)
C:Species: Gallus gallus (chicken)
C:Date: 06-Nov-1994 #sequence_revision 18-Nov-1994 #text_change 28-Apr-1995
C:Accession: A09480
R:Yom, A.W.; Beutchen, L.M.; Leone, J.W.; Heinrichson, R.L.; Deibel Jr., M.K.
Biochemistry 32, 12571-12576, 1993
A:Title: An avian p56c-binding domain of 17,000 daltons is isolated following limited
A:Reference number: A09480; M010:94072550
A:Accession: A09480
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-15 - VIM
A:Experimental source: Ovary
A:Keywords: cytochrome P-450; cytochrome P-450; cytochrome P-450; cytochrome P-450

```

```

Query Match 61.1% Score 22; DB 2; Length 15;
Best Local Similarity 80.0%; Pred. No. 1,5e-02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REAG 5
DB 2 REJTG 6

RESULT 6
SI:44598
endo 1,4 beta xylanase (E.C. 3.2.1.6) Thielavia terrestris (strain Z-56) (fragment)
C:Species: Thielavia terrestris
C:Date: 04-Mar-1997 #sequence_revision 04-Mar-1997 #text_change 26-Feb-1998
C:Accession: A44598
R:Yaqubli, M.
submitted to the Protein Sequence Database, March 1994
A:Reference number: A44593
A:Accession: A44598
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-30 - YAG.
A>Note: strain 255b
C:Superfamily: endo-1,4 beta xylanase, endo 1,4 beta xylanase homology
C:Keywords: glycosidase; hydrolase

Query Match 61.1% Score 22; DB 2; Length 30;
Best Local Similarity 80.0%; Pred. No. 3e-02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 REAGR 6
DB 21 EGAGO 25

RESULT 7
SI:6633
c-ret protein - chicken (fragment)
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: S150633
R:Hannink, M.; Temin, H.M.
Oncogene 5, 1843-1850, 1990
A:Title: Structure and autoregulation of the c-ret promoter.
A:Reference number: S150633; M010:9113748
A:Accession: S150633
A>Status: preliminary; Translated from cDNA/EMBL/DBEM
A:Molecule type: DNA
A:Residues: 1-9 - HAN.
A:Cross references: EMBL: X56443, M010:963338, FIDB: CAA09822.1; FIDB:5044483

Query Match 58.4% Score 21; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GAGR 6
DB 6 GAGR 9

RESULT 8
SI:4745
glutathione transferase (E.C. 2.5.1.18) 1-2 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Oct-1987 #sequence_revision 25-Oct-1987 #text_change 18-Jun-1993
C:Accession: B24735
R:Mannervik, B.; Auld, P.; Guthrie, C.; Jonsson, B.; Tabler, M.K.; Warholme, M.; Jon
Prest, R.; Auld, S.; Auld, A. 82, 7292-7295, 1995
A:Title: Identification of three classes of cytosolic glutathione transferase common
A:Reference number: A24745; M010:86042634

```


A:Accession: R04747
 A:Molecule type: Protein
 A:Superfamily: 118 -NAR-
 C:Keywords: glutathione transferase

Query Match 58.3%; Score 21; DB 2; Length 18;
 Best Local Similarity 80.0%; Prod. No. 4.6e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAGRM 7
 1 111
 1b 11 GIGRM 15

RESULT 7
 SD044
 C:Species: Mus musculus (house mouse)
 C:Date: 28 Feb 1990 #sequence_revision 28 Feb 1990 #text_change 30-May-1997
 C:Accession: S05505
 A:Title: Genomic organization of the genes encoding mouse T-cell receptor alpha chain.
 A:Reference number: M01016296.342
 A:Accession: S05505
 A:Molecule type: DNA
 A:Residues: 1-20 -WTN-
 A:Cross reference: EM01X02859
 A:Note: This sequence was determined from the germ line gene
 C:Keywords: T cell receptor

Query Match 58.3%; Score 21; DB 2; Length 20;
 Best Local Similarity 66.7%; Prod. No. 4.9e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAGRM 7
 1 111
 1b 4 GAGRM 9

RESULT 10
 SD074
 A:Title: alpha synuclein, NAC - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 27 Oct 1994 #sequence_revision oct-Nov 1996 #text_change 21 Jan 2000
 C:Accession: S06746
 A:Title: Genomic organization of the alpha synuclein gene.
 A:Reference number: S06746
 A:Accession: S06746
 A:Molecule type: Protein
 A:Residues: 1-28 -TEN-

Query Match 58.3%; Score 21; DB 2; Length 28;
 Best Local Similarity 100.0%; Prod. No. 4.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GAGRM 7
 1 111
 1b 4 GAGRM 27

RESULT 11
 PX0069
 C:Species: Pseudomonas syringae
 C:Date: 05 Apr 1994 #sequence_revision 05 Apr 1994 #text_change 07-May 1995

C:Accession: PX0069
 C:Species: Pseudomonas syringae
 C:Date: 114, 940 945, 1994
 A:Title: An inducible NADP-dependent D-phenylserine dehydrogenase from Pseudomonas S
 A:Reference number: PX0069; M010186519
 A:Accession: PX0069
 A:Molecule type: Protein
 A:Residues: 1-28 -PAC-
 A:Comment: This enzyme consists of two identical subunits. It catalyzes the oxidation
 C:Keywords: NADP; oxidoreductase
 F:13-16/Region: NADP binding status predicted

Query Match 58.3%; Score 21; DB 2; Length 28;
 Best Local Similarity 80.0%; Prod. No. 4.6e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAGRM 7
 1 111
 1b 13 GIGRM 17

RESULT 12
 PX0051
 C:Species: Fusarium sporotrichioides
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 06-Jun-1997
 C:Accession: PX0051
 A:Title: Genomic organization of the genes encoding mouse T-cell receptor alpha chain.
 A:Reference number: PX0051
 A:Accession: PX0051
 A:Molecule type: Protein
 A:Residues: 1-15 -CHG-

Query Match 55.4%; Score 20; DB 2; Length 15;
 Best Local Similarity 80.0%; Prod. No. 4.1e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 REGM 5
 1 111
 1b 8 REGM 12

RESULT 13
 SD044
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 21 Nov 1994 #sequence_revision oct-Nov 1996 #text_change 20-Jun-2000
 C:Accession: S04541; C44021
 A:Title: Genomic organization of the alpha synuclein gene.
 A:Reference number: S04541
 A:Accession: S04541
 A:Molecule type: DNA
 A:Residues: 1-17 -SCH-
 A:Cross reference: M010186519; M010186519; M010186519; M010186519
 A:Title: Genomic organization of the alpha synuclein gene.
 A:Reference number: S04541
 A:Accession: S04541
 A:Molecule type: Protein
 A:Residues: 1-17 -SCH-
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-17 -SCH-
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin

Query Match 55.4%; Score 20; DB 2; Length 17;
 Best Local Similarity 80.0%; Prod. No. 4.6e+02;

Matches 4: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GAGRM 7
 11 1
 14 6 GAGTM 14

RESULT 14

A33361
 NAME: regulated phosphoprotein, 21K - rat (fragment)
 Species: Rattus norvegicus (Norway rat)
 Release: oct-1997 #sequence_revision 08 Dec 1989 #text_change 25-oct-1996
 Accession: A33361
 Klemmings Jr., H.C.; Girault, J.A.; Williams, K.R.; LoPresti, M.B.; Greengard, P.
 J. Biol. Chem. 264, 7726-7734, 1989
 Affili: Appa-21, a cyclic AMP-regulated phosphoprotein (Mr 21,000) enriched in dephosphorylated studies of its phosphorylation in vitro.
 Reference number: A33361; MIM:89214228
 Accession: A33361
 Status: preliminary
 Molecular type: protein
 Residues: 1-19 CHRM
 Keywords: phosphoprotein

Query Match 55.6%; Score 20; DB 2; Length 19;

Best local Similarity 50.0%; Pred. No. 5,1002;
 Matches 4: Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 REGAGR 6
 11 1
 14 11 KSGAGK 16

RESULT 15

S14161
 Probable ribonucleoside transferase, 51K protein
 Name: Alternative names: 51K protein
 Species: Saccharopolyspora erythraea
 Release: Dec-Mar-1997 #sequence_revision 10-oct-1997 #text_change 21 Nov-1998
 Accession: S14161
 Kistritzky, P.; Green, R.; Packman, L.C.; Rawlings, R.J.; Staunton, J.; Leadlay, P.F.
 Eur. J. Biochem. 195, 823-830, 1991
 Affili: An acyl-carrier-protein - thioesterase domain from the 6-deoxyerythronolide B synthase.

Reference number: S14091; MIM:91154424

Accession: S14161
 Molecular type: protein
 Residues: 1-20 -CAF
 Genet loc:
 Gene: eryC11
 Function:
 Description: involved in desosamine biosynthesis
 Keywords: antibiotic biosynthesis; intramolecular oxidoreductase; isomerase

Query Match 55.6%; Score 20; DB 2; Length 20;

Best local Similarity 66.7%; Pred. No. 5,4002;
 Matches 4: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 REGAGR 6
 11 1
 14 6 RAGLR 11

Search completed: May 1, 2001, 11:26:20
 Job time: 125 sec

Genome version 4.5
Copyright (c) 1993-2000 Compugen Ltd.

Old protein: protein search against sw model

Run on: May 1, 2001, 11:24:13 : Search time 16.41 seconds
(without alignments)

14,977 million cell updates/sec

Input: US 09 341 829a 5 copy 142_148
Protein source: 6
Sequence: 1 REFSEQM 1

Scoring table: BLAST006Z
Gapop 16.00 : Gapext 0.5

Searches: 9446 seqs, 4255486 residues

Total number of hits satisfying chosen parameters: 1767

Minimum DB seq length: 6
Maximum DB seq length: 46

Post processor: Minimum Match, 6%
Maximum Match, 100%
Listed 1180 45 summaries

Left database: SWISSPROT 49:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 24 | 63.9 | 26 | 1 RL16_BACST | P23410 bacillus st |
| 2 | 20 | 55.6 | 16 | 1 ORF3_341D | P81420 solanum tub |
| 3 | 20 | 55.6 | 16 | 1 HPA2_C10FA | P81443 clostridium |
| 4 | 20 | 55.6 | 24 | 1 HPA1_PAB11 | P14571 erythrolacus |
| 5 | 20 | 55.6 | 26 | 1 E417_MAE03 | P11817 human adeno |
| 6 | 20 | 55.6 | 28 | 1 EF05_SL00N | P80715 streptococ |
| 7 | 19 | 52.8 | 14 | 1 SAR2_A6660 | P11750 arabidops |
| 8 | 19 | 52.8 | 15 | 1 F1RA_A0461 | P12801 anas platyr |
| 9 | 18 | 50.0 | 10 | 1 C9X0_FAT | P80432 rattus norv |
| 10 | 18 | 50.0 | 10 | 1 C9X0_H0906 | P80982 thomops obe |
| 11 | 18 | 50.0 | 24 | 1 KAD_BW111 | P55140 bacillus li |
| 12 | 18 | 50.0 | 25 | 1 NE01_C10CK | P44963 gallus gall |
| 13 | 17 | 47.2 | 14 | 1 IAT1_P00ME | P01470 tremella mo |
| 14 | 17 | 47.2 | 14 | 1 F1RA_B00SE | P14452 equus cabal |
| 15 | 17 | 47.2 | 15 | 1 F1RA_SYN7A | P14463 syncerus ca |
| 16 | 17 | 47.2 | 16 | 1 F1RA_C00ST | P14535 ceratotheri |
| 17 | 17 | 47.2 | 16 | 1 F1RA_B00AS | P14449 capus asini |
| 18 | 17 | 47.2 | 16 | 1 F1RA_F00CA | P14150 tetra melle |
| 19 | 17 | 47.2 | 16 | 1 F1RA_H00LA | P14453 hylobates l |
| 20 | 17 | 47.2 | 16 | 1 F1RA_M00TH | P12803 macaca fusc |
| 21 | 17 | 47.2 | 16 | 1 F1RA_M00LE | P14455 mandillus |
| 22 | 17 | 47.2 | 16 | 1 F1RA_G00HE | P14459 odonocollus |
| 23 | 17 | 47.2 | 16 | 1 F1RA_T00PE | P14536 tapirus ter |
| 24 | 17 | 47.2 | 17 | 1 F1RA_P00G | P14460 sus scrofa |
| 25 | 17 | 47.2 | 18 | 1 F1RA_C000R | P14444 camelus dro |
| 26 | 17 | 47.2 | 18 | 1 F1RA_L00ML | P14454 lama glama |
| 27 | 17 | 47.2 | 19 | 1 F1RA_B00S0 | P14441 boson bonas |
| 28 | 17 | 47.2 | 19 | 1 F1RA_B00B0 | P14442 bubalus bub |
| 29 | 17 | 47.2 | 19 | 1 F1RA_C00B0 | P14446 cervus elap |
| 30 | 17 | 47.2 | 19 | 1 F1RA_C00B1 | P14447 cervus elap |
| 31 | 17 | 47.2 | 19 | 1 F1RA_M00M0 | P14457 monticola m |
| 32 | 17 | 47.2 | 19 | 1 F1RA_S00B0 | P14451 ovis aries |
| 33 | 17 | 47.2 | 20 | 1 C00P_V00CH | P80406 verticillu |

ALIGNMENTS

RESULT 1

RL16_BACST STANDARD: P81: 26 AA.
AC P23410:
DT 01-NOV-1991 (Rel., 20, Created)
DT 01-NOV-1991 (Rel., 20, Last sequence update)
DT 01-A03-1992 (Rel., 23, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L16 (FRAGMENT).
GN RPLP.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI TaxID=1422;
RN 111
RP SEQUENCE FROM N.A.
RX MEDLINE 91025633; PubMed 2222862;
FA Kikuchi W., Hatakeyama T., Kimura M.:
RT "Nucleotide sequences of Bacillus stearothermophilus ribosomal
protein genes: part of the ribosomal S10 operon."
RL Biol. Chem. Hoppe-Seyler 371:641-646(1990).
CC 1- FUNCTION: THIS PROTEIN BINDS PRETTY TO 23S RIBOSOMAL RNA AND IS
LOCATED AT THE A-SITE OF THE PEPTIDYLTRANSFERASE CENTER.
CC 1- SIMILARITY: BELONGS TO THE L16P FAMILY OF RIBOSOMAL PROTEINS.

THIS SWISS PROTEIN ENTRY IS COPYRIGHTED. IT IS PRODUCED THROUGH A COLLABORATION
BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION
AT THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
ENTITIES REQUIRES A LICENSE AGREEMENT. IF YOU ARE A RESEARCHER, YOU
OR SEND AN EMAIL TO LICENSE@SWISS-PROT.CH.

CC EMBL: X54994; CAA08741.1;
CC PIR: S10614; S10614.
CC InterPro: IPR000114;
CC Pfam: PF00252; Ribosomal_L16; 1.
CC ProSITE: PS00346; P10750-AL-116; 1; PARTIAL.
CC PROSITE: PS00701; RIBOSOMAL_L16; 2; PARTIAL.
CC Ribosomal protein; tRNA binding.
CC NON_TER 26 26
CC SEQUENCE: 26 AA; 4225 MW; 54555.667/141638 CR604;

Query Match 63.9%; Score 24; Len 16; Length 26;
Best local similarity 71.4%; Pred. No. 82;
Matches 5; Conservation 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 REGARM 7

11 REGARM 17

RESULT 2

QY3_S01D0 STANDARD: P81: 15 AA.
ID QY3_S01D0
AC P81420:

```

SQ      SEQUENCE: 16 AA; 1570 MW;  E3AB430E56575785 C16%4;
      1-17 Multi; 55.6%; Score: 20; DR 1; Length 16;
      1-8 Multi; Similarity: 33.3%; Pct. Id: 2.26-62;
      Matches: 4; Conservative: 0; Mismatches: 1; Indels: 0; Gaps: 0;

QY      3 GAGRM 7
      ||||
DB      8 GAGTM 12

RESULT  4
1B      HPTA_RAB11
AC      P14571;
DT      01-JAN-1990 (Ref. 14, Created)
DT      01-JAN-1990 (Ref. 14, Last sequence update)
DT      01-APR-1991 (Ref. 19, Last annotation update)
DE      HEPATOPOETIN A LIGHT CHAIN (HPTA) (F8AC0HMT).
OS      Oryctolagus cuniculus (Rabbit).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OJ      Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX      NCBI_TaxID:9986;
RN      [1]
KI      SEQUENCE.
RE      MEGALINE:89492048; PubMed 2528955;
RA      Zarnegar K., Mada S., Dohild J., Michalopoulos G.;
RT      "NH2-terminal amino acid sequence of rabbit hepatopoietin A, a
ET      heparin binding polypeptide growth factor for hepatocytes."
RC      Biochem. Biophys. Res. Commun. 163:1370-1376 (1989).
CC      FUNCTION: HPTA IS AN ACIDIC HEPARIN BINDING GROWTH FACTOR FOR
      HEPATOCYTES.
CC      1- SUBUNIT: HETERODIMER OF AN HEAVY AND A LIGHT CHAIN LINKED BY
      DI SULFIDE BOND(S).
DR      PIR: A3242; A3292.
KW      Growth factor; Heparin-binding.
FT      RN_TER 24 24
SQ      SEQUENCE 24 AA; 2719 MW;  E7889F9B807252FE C16%4;

```

```

Query Match 55.8% Score 20; DB 1; Length 24;
Best Local Similarity 57.1%;
Matches 4; conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RECAGRM 7
   I I I I
DB 8 RNVGRM 14

RESULT 5
L417_ADE04
ID E417_ADE04 STANDARD; PRI: 26 AA.
AC P11817;
DT 01-OCT-1989 (Rel. 12, Created)
DI 01-OCT-1989 (Rel. 12, last sequence update)
DT 01-NOV-1997 (Rel. 45, last annotation update)
DE PROBABLE EARLY E4 17 KDA PROTEIN (FRAGMENT).
OS Homo adenovirus type 3.
OR Virus-like, distal, virulent, no RNA stage. Adenoviridae: Mastadenovirus.
GX NBI 14x10 45659;
KN 111
EF SEQUENCE FROM N.A.
EA MEDLINE 95100162; PubMed 2992041;
FA 214666; Abstract 15; Petersson U;
FT "Adenovirus 3-like polypeptide gives implications for the structure
   of the fiber protein".
RL 3. Virid. 53:572-676(1985).
CC
CC THE SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```



```

Query Match: 52.98; Score 19; DB 1; Length 15;
Best Local Similarity: 60.08; Pred. No. 3,6+02;
Matches: 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 REGAG 5
ID 1 1
DB 9 REGAG 13

RESULT 9
CXX0 RAI STANDARD; PRT: 10 AA.
AC P804.47
DI 01-NOV-1995 (Rel. 42; Created)
DI 01-NOV-1995 (Rel. 42; Last sequence update)
DI 30-MAY-2000 (Rel. 49; Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE VIIC (B*1.9.3.1) (V111A) (FRAGMENT).
GN CXX7C OR CXX7C1
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI:txid:10116;
KN 111
RF SEQUENCE.
RC STRAIN WISTAR, ISSUES Liver, and Heart.
RX MEDLINE 9524729; PubMed 7681462.
KA Schuster H., Boppe H., Hainke W., Perle H., von Dargatzis J.
KA "Cytochrome oxidase is encoded by 70 loci in Drosophila and 41
R1 amino-terminal sequences suggest identity of the total heart and the
R1 adult liver isoform".
RL Eur. J. Biochem. 230:225-241(1995).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +
CC 4 FERROCYTOCHROME C.
KW oxidoreductase; Mitochondrion.
FT NON_TER 10 10
SQ SEQUENCE 10 AA: 1117 MW: 1260767687B1R0B CRG64;

Query Match: 50.08; Score 18; DB 1; Length 10;
Best Local Similarity: 60.08; Pred. No. 3,6+02;
Matches: 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EYAR 6
ID 1 1
DB 6 EYAR 9

RESULT 10
CXX0 RHB04 STANDARD; PRT: 10 AA.
DI 01-NOV-1997 (Rel. 45; Created)
DI 01-NOV-1997 (Rel. 45; Last sequence update)
DI 01-NOV-1997 (Rel. 45; Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE VIIC (B*1.9.3.1) (FRAGMENT).
OS Thunnus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Scombriformes; Perciformes; Scombroidei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
CX NCBI:txid:2241;
KN 111
RF SEQUENCE.
RC ISSUE Heart, and Liver.
RX MEDLINE 97454291; PubMed:9310366;
KA Arnold S., Lee L., Kim M., Song E., Lindner D., Lottspeich F.,
KA Kadaneub, B.
R1 The subunit structure of cytochrome c oxidase from tuna heart and

```

```

Liver."
RL Eur. J. Biochem. 248:99-104(1997).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +
CC 4 FERROCYTOCHROME C.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE,
CC oxidoreductase; Inner membrane; Mitochondrion.
FT NON_TER 10 10
SQ SEQUENCE 10 AA: 1059 MW: 1260767687B1R0B CRG64;

Query Match: 50.08; Score 18; DB 1; Length 10;
Best Local Similarity: 60.08; Pred. No. 3,6+02;
Matches: 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EGAG 6
ID 1 1
DB 5 EGAG 9

RESULT 11
KAL0 BACLI STANDARD; PRT: 24 AA.
AC P35140;
DI 01-FEB-1994 (Rel. 28; Created)
DI 01-FEB-1994 (Rel. 28; Last sequence update)
DI 15-JUL-1999 (Rel. 36; Last annotation update)
DE ACETYLATE KINASE (E*2.7.4.3) (ACT AKG TRANSFER-SUBCLASE) (FRAGMENT).
GN ADK.
OS Bacillus licheniformis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
CX NCBI:txid:1402;
KN 111
RF SEQUENCE FROM N.A.
RX MEDLINE 93062802; PubMed 1445226;
KA Tschander S., Driessen A.J.M., Freudl R.
RT "Cloning and molecular characterization of the secY genes from
RT Bacillus licheniformis and Staphylococcus carnosus: comparative
RT analysis of nine members of the SecY family."
RL Mol. Gen. Genet. 245:147-152(1992).
CC -!- FUNCTION: THIS SMALL GTP-BINDING ENZYME IS ESSENTIAL FOR
CC MAINTENANCE AND CELL GROWTH.
CC -!- CATALYTIC ACTIVITY: ATP + AMP -> ADP + ADP.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE ACETYLATE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. See http://www.sib.ac.ch/eng/
CC or send an email to license@sib.ac.ch)
CC EMBL: X70087; -; NOT_ANNOTATED CDS.
DR PIR: S34405; S34405.
DR BSSP: P27142; 1210.
DR InterPro: IPR000850;
DR Pfam: PF09409; adenylyl-kinase; 1.
DR PROSITE: PS00113; ADENYLYL_KINASE; PARTIAL.
KW transferase; Kinase; ATP-binding.
FT NON_TER 24 24
SQ SEQUENCE 24 AA: 2563 MW: A6501A5F1010F21 CRG64;

Query Match: 50.08; Score 18; DB 1; Length 24;
Best Local Similarity: 75.09; Pred. No. 8,6+02;
Matches: 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```


EX HEADLINE: 67209145; PubMed 6037211;
 FA Description R.F.; Schubert D.; Schwartz S.A.;
 FI "Amino acid sequence studies on artiodactyl fibrinopeptides. I.
 FI Dromedary camel, mule deer, and cape buffalo.";
 FL Arch. Biochem. Biophys. 118:456-467(1967).
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THIS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 KW blood coagulation; Plasma.
 FI N_TERM 15
 SL SEQUENCE 15 AA. 1480 MW. 4E996A510B410C6 CRC64;

Query Match 47.2%; Score 17; DB 1; Length 15;
 Best Local Similarity 75.0%; Pred. No. 8.5e+02;
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QV 2 E5A3 5
 DB 10 E5A3 14

Search completed: May 1, 2001, 11:25:50
 Job time: 100 sec

GenCore version 4.5
Copyright (c) 1994-2000 Compugen Ltd.

1M protein protein search, using SW model

Run out: May 1, 2001, 11:25:52 : Search time 40.59 seconds
(without alignments)
22.533 M11132.rcl, updates/sec

Hit list: US 09 341 829A-5_copy 164 179
Percent score: 92
Sequence: 1 KAVGAPDPKHKVSEKQ 16

Scoring table: BL08062
Gap: 10.0 x GapExt 0.5

Search: 30/29 seqs, 5/163235 residues 177502
Total number of hits satisfying chosen parameters:

Minimum hit seq length: 5
Maximum hit seq length: 30

Post processing: Maximum Match 98
Listing first 45 summaries

| Database: 1 A Jonesseq 0401:* | | | |
|-------------------------------|-----|-----|-----|
| 1 | 2 | 3 | 4 |
| 1 | 111 | 111 | 111 |
| 2 | 111 | 111 | 111 |
| 3 | 111 | 111 | 111 |
| 4 | 111 | 111 | 111 |
| 5 | 111 | 111 | 111 |
| 6 | 111 | 111 | 111 |
| 7 | 111 | 111 | 111 |
| 8 | 111 | 111 | 111 |
| 9 | 111 | 111 | 111 |
| 10 | 111 | 111 | 111 |
| 11 | 111 | 111 | 111 |
| 12 | 111 | 111 | 111 |
| 13 | 111 | 111 | 111 |
| 14 | 111 | 111 | 111 |
| 15 | 111 | 111 | 111 |
| 16 | 111 | 111 | 111 |
| 17 | 111 | 111 | 111 |
| 18 | 111 | 111 | 111 |
| 19 | 111 | 111 | 111 |
| 20 | 111 | 111 | 111 |
| 21 | 111 | 111 | 111 |
| 22 | 111 | 111 | 111 |

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 35 | 42.7 | 46 | 8 p70916 | Partial sequence o |
| 2 | 33 | 40.4 | 9 | 16 R07278 | Plasmodium falcipa |
| 3 | 33 | 40.2 | 9 | 20 Y46560 | Immunogenic peptid |
| 4 | 32 | 39.9 | 13 | 16 W21426 | Glucuron precursor |
| 5 | 31 | 37.8 | 12 | 18 W25294 | Peptide clone BpD- |
| 6 | 31 | 37.8 | 30 | 20 Y60492 | Human normal bladd |
| 7 | 30 | 36.6 | 13 | 16 R74800 | Human tumour necro |
| 8 | 30 | 36.6 | 19 | 16 R67677 | Mouse kappa/delta |
| 9 | 30 | 36.6 | 26 | 20 Y50937 | Human myometrium t |
| 10 | 30 | 36.6 | 22 | 7 P61403 | Peptide acting sel |
| 11 | 30 | 36.6 | 22 | 7 P61404 | Peptide acting sel |

| | | | | | | |
|----|------|------|----|----|--------|-----------------------|
| 12 | 30 | 36.6 | 22 | 7 | P61405 | Peptide acting sel |
| 13 | 30 | 36.6 | 40 | 21 | G38051 | Arabinoside thiaia |
| 14 | 29 | 35.4 | 9 | 20 | Y55614 | HLA binding pla 1 |
| 15 | 29 | 35.4 | 15 | 16 | R01715 | C. albicans enolase |
| 16 | 29 | 35.4 | 15 | 18 | W22567 | S. pneumoniae RS9-7 |
| 17 | 29 | 35.4 | 21 | 21 | W04535 | Bovine factor V ac |
| 18 | 29 | 35.4 | 25 | 17 | W04444 | HLV and bovine pep |
| 19 | 28.5 | 34.1 | 18 | 18 | W12709 | Anti-hepatitis pep |
| 20 | 28 | 34.1 | 9 | 20 | Y55514 | HLA binding pla 1 |
| 21 | 28 | 34.1 | 14 | 21 | W25222 | Helicobacter pylori |
| 22 | 28 | 34.1 | 20 | 12 | R11301 | Recombinant 175-19 |
| 23 | 28 | 34.1 | 20 | 16 | W13196 | Fragment of p53 f1 |
| 24 | 28 | 34.1 | 20 | 19 | W52347 | Human W601 immunog |
| 25 | 28 | 34.1 | 21 | 9 | P31560 | Sequence of thinn |
| 26 | 28 | 34.1 | 21 | 21 | W04536 | Mouse factor V act |
| 27 | 28 | 34.1 | 28 | 14 | R66513 | 132-39 anti-hep-18 |
| 28 | 28 | 34.1 | 28 | 15 | R56750 | Random peptide #1 |
| 29 | 28 | 34.1 | 28 | 17 | R91498 | 132-39 monoclonal |
| 30 | 28 | 34.1 | 28 | 18 | W52383 | Anti-body 132-39 ep |
| 31 | 27 | 32.9 | 9 | 15 | R62770 | Anti-hepatitis fo vir |
| 32 | 27 | 32.9 | 9 | 20 | Y55513 | HLA binding pla 1 |
| 33 | 27 | 32.9 | 9 | 20 | Y55613 | HLA binding pla 1 |
| 34 | 27 | 32.9 | 12 | 14 | R60954 | Peptide Ia from 17 |
| 35 | 27 | 32.9 | 14 | 16 | R93516 | Class II c5 GS1 YE |
| 36 | 27 | 32.9 | 14 | 16 | R93489 | Class II c5 GS1 YE |
| 37 | 27 | 32.9 | 15 | 10 | P90206 | Anti-hepatitis pep |
| 38 | 27 | 32.9 | 18 | 19 | Y20305 | Human microtubule |
| 39 | 27 | 32.9 | 19 | 9 | P82531 | Region of P. falcip |
| 40 | 27 | 32.9 | 19 | 14 | R60956 | Peptide Ia from 17 |
| 41 | 27 | 32.9 | 20 | 16 | R71653 | c5 Region 1 subset 1 |
| 42 | 27 | 32.9 | 20 | 20 | W04535 | Human secreted pro |
| 43 | 27 | 32.9 | 21 | 20 | W04536 | Human secreted pro |
| 44 | 27 | 32.9 | 21 | 20 | W04537 | Human secreted pro |
| 45 | 27 | 32.9 | 25 | 22 | W04672 | Human secreted pro |

Accession

| RESULT | 1 |
|--------|--|
| P70916 | |
| 1 | Partial sequence of amino terminal of parathyroid hormone (PTH). |
| 2 | Anti-hepatitis fo virus (anti-HFV) monoclonal antibody. |
| 3 | Host-strain rejection. |
| 4 | Rattus rattus. |
| 5 | Key |
| 6 | Misc-difference |
| 7 | Key |
| 8 | Misc-difference |
| 9 | Key |
| 10 | Misc-difference |
| 11 | Key |
| 12 | Misc-difference |
| 13 | Key |
| 14 | Misc-difference |
| 15 | Key |
| 16 | Misc-difference |
| 17 | Key |
| 18 | Misc-difference |
| 19 | Key |
| 20 | Misc-difference |
| 21 | Key |
| 22 | Misc-difference |
| 23 | Key |
| 24 | Misc-difference |
| 25 | Key |
| 26 | Misc-difference |
| 27 | Key |
| 28 | Misc-difference |
| 29 | Key |
| 30 | Misc-difference |
| 31 | Key |
| 32 | Misc-difference |
| 33 | Key |
| 34 | Misc-difference |
| 35 | Key |
| 36 | Misc-difference |
| 37 | Key |
| 38 | Misc-difference |
| 39 | Key |
| 40 | Misc-difference |
| 41 | Key |
| 42 | Misc-difference |
| 43 | Key |
| 44 | Misc-difference |
| 45 | Key |

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 35 | 42.7 | 46 | 8 p70916 | Partial sequence o |
| 2 | 33 | 40.4 | 9 | 16 R07278 | Plasmodium falcipa |
| 3 | 33 | 40.2 | 9 | 20 Y46560 | Immunogenic peptid |
| 4 | 32 | 39.9 | 13 | 16 W21426 | Glucuron precursor |
| 5 | 31 | 37.8 | 12 | 18 W25294 | Peptide clone BpD- |
| 6 | 31 | 37.8 | 30 | 20 Y60492 | Human normal bladd |
| 7 | 30 | 36.6 | 13 | 16 R74800 | Human tumour necro |
| 8 | 30 | 36.6 | 19 | 16 R67677 | Mouse kappa/delta |
| 9 | 30 | 36.6 | 26 | 20 Y50937 | Human myometrium t |
| 10 | 30 | 36.6 | 22 | 7 P61403 | Peptide acting sel |
| 11 | 30 | 36.6 | 22 | 7 P61404 | Peptide acting sel |

FI New peptide parathyromisin alpha - having immunomodulatory activity and useful for treating autoimmune disease or patients receiving organ transplants

XX claim 1; column 6; 7pp; English.

XX p70916 is the partial N-terminal SQ of a claimed peptide (PTA) of about 100 AAs having a pI of about 4.15; PTA acts as a blocker or modulator to the immuno-enhancing activity exhibited by prothymosin alpha. Deems is 1100-005694-0417, given in ex. 1, 2, or 3, c.

XX Sequence 30 AA;

SQ Sequence 30 AA;

Query Match 42.7%; Score 35; PR 8; Length 40;

Best Local Similarity 50.0%; Pred. No. 16;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Caps 0;

CV 2 AR6LRPPKPVKSPK 15

1413 11115

15 AP104KKKPKKK 26

RESULT 2

867278

XX 867278 standard; peptide; 9 AA.

XX 867278;

XX 16-MAY-1996 (first entry)

XX Plasmodium falciparum HLA-B8 restricted epitope cp43.

XX circumsporozoite protein; human leukocyte antigen; HLA, class I, HLA-B8; epitope; malaria; vaccine; CTL induction;

XX cytotoxic T lymphocyte.

XX Plasmodium falciparum.

XX W09506982-A2.

XX 12-01-1995.

XX S1 MAP-1995; 95W-GB0347.

XX S2 MAR-1994; 94GB-00064922.

XX (ISIS-1) ISIS INNOVATION LTD

XX Above M. Allison (TM) p01 AV*, Patrick A. Fitzpatrick B.

XX White BC;

XX WPI: 1995-356584/46.

XX Plasmodium falciparum peptide(s) - useful in vaccine compositions for immunising against malaria

XX claim 1; Page 18; 23pp; English.

XX cytotoxic T lymphocytes from malaria exposed Gambian individuals with HLA-B8 showed significant lysis of target cells pre-pulsed with a pool of the four peptides cp43, cp44, cp45 and cp46 (all derived from Plasmodium falciparum (Fowcett-Pratt) strain 3D7, one of the four peptides, i.e. cp43, bound to HLA-B8 in the HLA assembly assay; identifying it as an HLA-B8 restricted epitope.

XX the cp peptides will be useful in a malaria vaccine.

XX Sequence 9 AA;

Query Match

Best Local Similarity 40.2%; Score 33; PR 16; Length 9;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Caps 0;

CV 5 LSTPPKHK 11

11 1111

11 KKPKKK 7

RESULT 3

Y46560

XX Y46560 standard; Peptide; 9 AA.

XX Y46560;

XX 01-DEC-1999 (first entry)

XX immunogenic peptide having a human leukocyte antigen binding motif #1171.

XX human leukocyte antigen, binding, immunogenic, glycoprotein; MHC; HLA;

XX immune response; T cell activation; major histocompatibility complex;

XX cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;

XX prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;

XX vaccine; immunisation.

XX Synthetic.

XX Homo sapiens.

XX W09945954-A1.

XX 16 SEP 1999.

XX 14-MAR-1998; 98WO-US05039.

XX 14-MAR-1998; 98WO-US05039.

XX (EP18) EP180396 IPC.

XX Sette A, Kalle ET, Sidney J, Cellis E, Grey HM, Berlewood S;

WPI: 1999-551214/46.

XX New immunogenic peptides with HLA binding motif, useful in treatment and diagnosis of cancers and viral diseases

XX claim 1; Page 76; 150pp; English.

XX Y45390 to Y48214 represent specifically claimed immunogenic peptides having a human major histocompatibility complex (MHC) Class I (also known as human leukocyte antigen (HLA)) binding motif. The immunogenic peptides can bind to a specific HLA allele (i.e. HLA subtypes HLA-A2.1, A1, A3.2 or A24.1 or HLA B or C) and induce a cytotoxic T cell response against the antigen from which the peptide is derived.

XX cytotoxic T lymphocytes (CTLs) which destroy antigen bearing cells are normally induced by an antigen in the form of a peptide fragment bound to a HLA molecule, rather than the intact foreign antigen itself, and are particularly important in tumour rejection and in fighting viral infections. The peptides are therefore useful therapeutically to treat or prevent viral infections and cancers in mammals (especially humans) e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.

XX they can be administered as vaccines to elicit an immune response in individuals susceptible or otherwise at risk of viral infection or cancer, or used to treat chronic or acute conditions. They are also useful diagnostically, and can be used to induce a cytotoxic T cell response, by contacting a cytotoxic T cell with the peptide e.g. to induce CTLs to kill for infection back into a patient. The polynucleotides encoding the immunogenic peptides are also useful therapeutically and for immunisation as above.

XX Sequence 9 AA;

Query Match

Best Local Similarity 40.2%; Score 33; PR 20; Length 9;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Caps 0;

```

Q7 5. LKPKHK 11
   1 LKPK
   2 LKPKHK
Db 2 LKPKHK 6
Query Match 59.00% Score 327 Db 16 Length 13
Best local similarity 66.7% Pos. No. 22
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0

PEP001 4
W21420
Db W21420 standard; peptide: 13 AA
XX
XX
XX W21420
L1 20 JUL 1997 (first entry)
XX
XX Glucuronidase, precursor-derived signal oligopeptide #25.
XX
XX Hydrophilic; signal oligopeptide; hydrophilicity maxima; varicose; HIV;
XX competitive inhibitor; feedback regulator; synthesis; gastrin precursor;
XX chaperone; poliovirus; fatty acid synthase; plasminogen activator inhibitor 1;
XX hydroxymethylglutaryl coenzyme A reductase; glucagon precursor; ribosyl;
XX quonilidein precursor; plasminogen activator inhibitor 2; prorenin;
XX Alzheimer amyloid A; corticotropin releasing factor binding protein;
XX hepatitis B; herpes virus 1 glycoprotein B; HSV1; human; OMV5;
XX herpes virus 2 glycoprotein B; HSV2; collagenase; apolipoprotein A;
XX tripeptidyl peptidase; membrane protein; MMPA; islet amyloid polypeptide;
XX fibrinase; MMP1; schistosoma elastase precursor; schistosomin;
XX hepatitis delta antigen; rev protein; HIV; HIV; angiotensinogen.
XX
XX Bone sponges.
XX
XX K% Location/Qualifiers
XX
XX Mismatch difference 10
XX /note "undefined amino acid"
XX
XX W06510568 A1
XX
XX 20 JUL 1995
XX
XX 12 JAN 1995; 96WS 0300475.
XX
XX 14 JAN 1994; 94TS 0162246.
XX
XX (RAH/2) RAH R.
XX
XX Rath R.
XX
XX WPI: 1995 2609574.
XX
XX Identifying signal oligopeptides(s) in protein sequence(s) - shown as
XX regions of max. hydrophilicity, used in modulating communication
XX between proteins(s)
XX
XX Claim 5; Page 42; 88pp; English.
XX
XX The sequences given in W21420 represent hydrophilic signal oligo-
XX peptides. These signal oligopeptides are localised on the surface
XX of the protein and are represented by the hydrophilicity maxima of
XX the protein. These peptides are enriched in charged amino acids
XX associated with neutral spacer amino acids. The specific signal
XX character of these oligopeptides is determined by a characteristic
XX combination of conformation and charge within the signal sequence.
XX These oligopeptides may be used as vaccines in the treatment of
XX human disease, as competitive inhibitors to prevent or reduce the
XX metabolic action or interaction of a selected protein by blocking
XX its specific signal sequences, or as therapeutic agents to function
XX as feedback regulators to reduce synthesis rate of a selected protein.
XX These peptides may be modified by omitting one or more amino acids at
XX the N- and/or C-terminal, by substituting one or more amino acids
XX without consideration of charge and polarity, by substituting one or
XX more amino acids with amino acid residues with similar charge and/or
XX polarity, by omitting one or more amino acids or a combination of these.
XX
XX Sequence 13 AA:

```

```

Query Match 49.00% Score 327 Db 16 Length 13
Best local similarity 66.7% Pos. No. 22
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0

Q7 1 KARLGRUPK 9
   1111111
Db 1 KARLGRUPK 9
Query Match 59.00% Score 327 Db 16 Length 13
Best local similarity 66.7% Pos. No. 22
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0

RESULT 5
W25294
Db W25294 standard; peptide: 12 AA
XX
XX W25294
L1 14 OCT 1997 (first entry)
XX
XX Peptide clone Bp0-7 specific for monoclonal antibody b42.09.
XX
XX Monoclonal antibody b42.09; Lact; headpiece domain; DNA binding protein;
XX random peptide library; receptor ligand; dimer; fusion protein; epitope;
XX antibody.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX
XX Region 2.18
XX /note "b42.09 antibody epitope"
XX
XX W09640987 A1
XX
XX 19 DEC 1996
XX
XX 07 JUN 1996; 96WS 0509409.
XX
XX 26 OCT 1995; 95WS 0648546.
XX
XX 07 JUN 1995; 95WS 0484090
XX
XX (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
XX Cell M5; Gates CM; Miller JF; Schatz PJ; Stromer WB.
XX
XX WPI: 1997 087065708.
XX
XX Random peptide library and affinity enrichment methods for screening
XX it - useful to identify peptide(s) that bind receptor mol(s) of
XX interest, useful for therapeutic, diagnostic and related purposes
XX
XX Example 5; Fig 8; 14pp; English.
XX
XX W25289 W25289 represent b42.09 monoclonal antibody specific peptides.
XX These sequences were isolated by a method of the invention to isolate a
XX DNA binding protein, or a peptide with specific affinity for a receptor.
XX The method comprises providing a recombinant BNA vector encoding a
XX peptide having specific affinity for a receptor. A library of
XX oligonucleotides encoding different potential DNA binding proteins is
XX inserted in frame into the vector to create a fusion protein library.
XX Host cells are transformed, and cultured to express the fusion protein.
XX If a fusion protein comprises a potential DNA binding protein with
XX affinity for the vector, the fusion protein binds to the vector to form a
XX complex. The host cells are lysed to isolate the complexes which are
XX contacted with a receptor to induce peptide binding to the receptor. The
XX random peptide library and the methods for screening it can be used to
XX identify peptides that bind receptor molecules of interest. The peptides
XX can be used for therapeutic, diagnostic and related purposes, e.g. to
XX bind the receptor, or an analogue, and so inhibit or promote the activity
XX of the receptor. The method of affinity enrichment allows a very large
XX library of peptides to be screened, and by identifying the peptide de-
XX novo, the sequence or structure of the receptor need not be known.
XX
XX Sequence 12 AA:

```

Query Match 37.88; Score 31; DB 18; Length 12;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 6; Conservations 0; Mismatches 0; Indels 0; Gaps 0;

QV 10 HKVSEQ 15
 111111
 10 4 HKVSEQ 9

RESULT 6
 Y60492
 ID Y60492 standard; Protein: 30 AA.
 XX
 AC Y60492;
 DT 01-JAN-2000 (first entry)
 XX
 DE Human normal bladder tissue EST encoded protein 164.
 XX
 KW human bladder treatment; EST encoded sequence; cytotoxicity;
 KW cancer; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PR 01-NOV-1999;
 XX
 PR 21-APR-1998; 98DE-1018620.
 XX
 PR 21-APR-1998; 98DE-1018620.
 XX
 FA (META-) METACIN CHEMOPREVENTIVE REB.
 XX
 RT Rosenthal A, Specht T, Hinzmann R, Schmitt A, Pillarsky G, Bahl R;
 XX
 DB WPI; 1999-002416/92.
 DB N-PSDB; 242208.
 XX
 XX New polypeptides and their nucleic acids, useful for treatment of
 PT Bladder tumour and identification of therapeutic agents
 XX
 PS Claim 24; Page 311; 866pp; German
 XX
 XX This invention describes novel polypeptide fragment sequences (I) and
 CC their encoding nucleic acids (II) which are highly expressed in normal
 CC bladder tissue and have cytostatic activity. (II) are used for
 CC treatment of (I) and to isolate complete genes. (I) are
 CC used to identify agents suitable for the treatment of bladder tumours, to
 CC directly treat this form of cancer (including expression from gene
 CC therapy vectors), or are used in a preparation for cancer treatment. (I)
 CC is also used for the generation of specific antibodies. (II) are
 CC identified by assembling ESTs (expressed sequence tags) from a particular
 CC tissue type before comparison of expression patterns. This allows a
 CC significantly longer fragment of the gene to be revealed, and therefore
 CC reduces the number of failures because of ESTs from different libraries
 CC representing different parts of the same unknown gene distorting the
 CC estimated frequency of occurrence in a particular tissue. Y6029-Y6094
 CC represent protein fragments encoded by the human normal bladder tissue
 CC cDNA library derived EST fragments represented in 242122-242248.
 XX
 SL Sequence 30 AA;

Query Match 37.88; Score 31; DB 20; Length 30;
 Best Local Similarity 62.5%; Pred. No. 82;
 Matches 6; Conservations 0; Mismatches 1; Indels 0; Gaps 0;

QV 6 RTHKRVK 13
 111111
 10 24 RTHKRVK 43

RESULT 7
 R74800
 ID R74800 standard; Protein: 17 AA.
 XX
 AC R74800;
 DT 09-JAN-1996 (first entry)
 XX
 DE Human tumour necrosis factor (hTNF) alpha (His 11).
 XX
 KW Interleukin-1 beta nonapeptide; IL-1 beta; fusion protein;
 KW human tumour necrosis factor alpha; hTNF alpha His 11; 163-171;
 KW irradiation protection; haematopoietic progenitor regeneration;
 KW immunostimulation; cancer cell growth reduction.
 XX
 OS Homo sapiens.
 XX
 PR Key Location/Qualifiers
 FT 1-13
 FT 1-13 "hTNF alpha His 11 residues 1-13
 FT 14-157 (residues 14-157 are absent)"
 XX
 PR W09512675 A1.
 XX
 PR 11-MAY-1995.
 XX
 PR 03-NOV-1994; 94WO-0812656.
 XX
 PR 04-NOV-1993; 94US-0147673.
 XX
 PA (CYTO) CYTOTOXICAL PHARM INC.
 PA (CYTO) CYTOTOXICAL PHARMACEUTICALS INC.
 XX
 PT Bolton AP, Siddhu RS;
 XX
 DB WPI; 1995-185778/24.
 DB N-PSDB; Q90266.
 XX
 XX New fusion protein for use in tumour therapy comprising a tumour
 PT necrosis factor molecule and interleukin 1 beta residues 153-171.
 XX
 PS Claim 4; Fig 3; 44pp; English.
 XX
 CC Q90266 encodes R74800 hTNF alpha His 11 (only residues 1-13 given),
 CC which was fused to the IL-1 beta residues 163-171 nonapeptide to
 CC produce a fusion protein. The fusion protein can be used to enhance
 CC protection against irradiation, and to accelerate the regeneration
 CC of haematopoietic precursors in animals treated with cytotoxic
 CC anti-cancer drugs and irradiation. In particular, they can be used to
 CC immunostimulate patients that have undergone irradiation therapy,
 CC they may also be used to selectively reduce the growth of cancer
 CC cells.
 XX
 SL Sequence 17 AA;

Query Match 36.64; Score 63; DB 16; Length 14;
 Best Local Similarity 84.3%; Pred. No. 51;
 Matches 5; Conservations 0; Mismatches 1; Indels 0; Gaps 0;

QV 6 RTHKRVK 11
 11111
 10 7 RTHKRVK 12

RESULT 8
 R67677
 ID R67677 standard; Peptide: 18 AA.
 XX
 AC R67677;
 XX

D1 18 AUG 1995 (first entry)
 XX Mouse kappa/delta opioid receptor second intracellular loop fragment.
 D2
 XX Mouse kappa/delta mu opioid receptor; brain; primer; PCR; amplify;
 KW transmembrane domain; somatostatin receptor; human; expression vector;
 KW truncated; chimera; assay; probe.
 D3 Synthesis.
 XX
 D4 W0424032 A.
 D5 06 DEC 1994.
 XX
 D6 20 MAY 1994; 94WO 0505747.
 XX
 D7 20 MAY 1994; 94OS 0066246.
 D8 30 JUL 1994; 94OS 0100694.
 D9 05 NOV 1994; 94OS 0147592.
 XX
 D10 (Aber) ABER DEV CHAP.
 XX
 D11 Bell GJ. Peptides L. Yasuda K.
 D12 WPI: 199 02004204
 XX
 D13 Peptide tubes and peptides derived from opioid receptor
 D14 polypeptides for use in therapeutic compositions and in
 D15 screening assays for useful drug substances.
 XX
 D16 (inverted source; Page 100; 40pp; English).
 XX
 D17 The amino acid sequence of the second intracellular loop of the mouse
 D18 kappa/delta opioid receptors. The sequence has similarity to the second
 D19 intracellular loop of the somatostatin receptor subtype 3STK1 (867676)
 D20 the mouse kappa and delta sequences are identical. The second and third
 D21 (94/674 C) loops of the opioid receptors are potentially involved in
 D22 interaction with G proteins in a signal mechanism. Novel chimeraic
 D23 receptors were constructed where the second or third intracellular loops
 D24 of the somatostatin receptor was exchanged with the second and/or third
 D25 loops of the kappa and/or delta mu opioid receptors. The role of
 D26 these loops could be determined by observing whether the somatostatin
 D27 receptors gained the ability to interact with G-proteins or if the opioid
 D28 receptors lost the ability to interact with G-proteins. The chimeraic
 D29 receptors were constructed by recombination and ligation of PCR amplified
 D30 fragments (See 975942-49 for primers). The chimeraic opioid receptors
 D31 thus produced are useful for the development of novel assays designed to
 D32 select or improve substances, capable of interacting with the opioid
 D33 receptor proteins, for use in diagnostics, drug design and therapeutic
 D34 applications.
 XX
 D35 Sequence: 20 AA;
 Query Match: 96.6%; Score 40; DB 16; Length 18;
 Best Local Similarity: 75.0%; Prod. No. 74;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 D36 1 KAROLIRP R
 11 1111
 D37 11 Kallirp 18
 RESULT 9
 Y9947
 D38 Y9947 standard; Protein: 20 AA.
 XX
 D39 Y9947;
 XX
 D40 28 JAN 2000 (first entry)
 XX
 D41 Human myometrium tumour EST encoded protein 17.
 D42
 XX

KW Myometrium; tumour; human; expressed sequence tag; EST; uterine myoma;
 KW treatment; carcinoma; cancer; gene therapy.
 D43 Homo sapiens.
 D44 DE19817947-A1.
 D45 28 SEP 1999.
 D46 17 APR 1998; 9803E 1017947.
 D47 17 APR 1998; 9803E 1017947.
 D48 (META) METAFGER GTS GENEEXPRESSURE MBL.
 D49 Rosenthal A, Specht T, Hinzmann K, Schmitt A, Florsky P, Dahl K.
 D50 WPI: 1999-602480/52.
 D51 N-PSDB: Z41980.
 D52
 D53 New nucleic acid sequences expressed in uterine myoma, and derived
 D54 polypeptides, for treatment of uterine carcinoma and identification of
 D55 therapeutic agents
 D56
 D57 Claim 23; Page 71; 86pp; German.
 D58
 D59 This invention describes novel polypeptide sequences (I). Fragments of
 D60 (I) fragments and their encoding nucleic acids (II) which are highly
 D61 expressed in human uterine myoma. (II) are used for recombinant
 D62 expression of (I) and to isolate complementary DNAs. (I) are used to
 D63 identify agents suitable for treatment of uterine carcinoma, to directly
 D64 treat this form of cancer (includng expression from gene therapy
 D65 vectors) and are used in a preparation for cancer treatment. (I) is also
 D66 used for the generation of specific antibodies. (II) are identified by
 D67 assembling ESTs (expressed sequence tags) from a particular tissue type
 D68 before comparison of expression patterns. This allows a significantly
 D69 longer fragment of the gene to be revealed and therefore reduces the
 D70 number of failures associated with the fact that ESTs from different
 D71 libraries may represent different parts of the same unknown gene.
 D72 Distorting the estimated frequency of occurrence in a particular tissue.
 D73 Y9921 Y9949 represent protein fragments encoded by the human
 D74 myometrium tumour cDNA library derived EST fragments represented in
 D75 Z41950 Z41980.
 D76
 D77 Sequence: 20 AA;
 Query Match: 96.6%; Score 40; DB 20; Length 20;
 Best Local Similarity: 50.0%; Prod. No. 80;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 D78 1 KAROLIRP R
 11 111111
 D79 6 Kallirp 19
 RESULT 10
 D80 P61403
 D81 P61403 standard; Protein: 22 AA.
 D82
 D83 P61403;
 D84
 D85 19 JUN 1991 (first entry)
 D86
 D87 Peptide acting selectively on the electric potential dependent sodium
 D88 channel.
 D89
 D90 Na channel; ATPase.
 D91
 D92 Cons. geographicals.
 D93
 D94 Key location/Qualifiers
 D95 Modified-site 6

FR 12-AUG-1999: 9908-0148341.
 FR 13-AUG-1999: 9908-0148365.
 FR 14-AUG-1999: 9908-0148684.
 FR 15-AUG-1999: 9908-0149368.
 FR 16-AUG-1999: 9908-0149175.
 FR 17-AUG-1999: 9908-0149426.
 FR 18-AUG-1999: 9908-0149722.
 FR 19-AUG-1999: 9908-0149723.
 FR 20-AUG-1999: 9908-0149929.
 FR 21-AUG-1999: 9908-0149902.
 FR 22-AUG-1999: 9908-0149930.
 FR 23-AUG-1999: 9908-0150566.
 FR 24-AUG-1999: 9908-0150884.
 FR 25-AUG-1999: 9908-0151065.
 FR 26-AUG-1999: 9908-0151066.
 FR 27-AUG-1999: 9908-0151080.
 FR 28-AUG-1999: 9908-0151080.
 FR 29-AUG-1999: 9908-0151303.
 FR 30-AUG-1999: 9908-0151303.
 FR 31-AUG-1999: 9908-0151340.
 FR 01-SEP-1999: 9908-0151340.
 FR 02-SEP-1999: 9908-0152364.
 FR 03-SEP-1999: 9908-0153070.
 FR 04-SEP-1999: 9908-0153758.
 FR 05-SEP-1999: 9908-0154018.
 FR 06-SEP-1999: 9908-0154039.
 FR 07-SEP-1999: 9908-0154779.
 FR 08-SEP-1999: 9908-0155139.
 FR 09-SEP-1999: 9908-0155139.
 FR 10-SEP-1999: 9908-0155139.
 FR 11-SEP-1999: 9908-0155139.
 FR 12-SEP-1999: 9908-0155139.
 FR 13-SEP-1999: 9908-0155139.
 FR 14-SEP-1999: 9908-0155139.
 FR 15-SEP-1999: 9908-0155139.
 FR 16-SEP-1999: 9908-0155139.
 FR 17-SEP-1999: 9908-0155139.
 FR 18-SEP-1999: 9908-0155139.
 FR 19-SEP-1999: 9908-0155139.
 FR 20-SEP-1999: 9908-0155139.
 FR 21-SEP-1999: 9908-0155139.
 FR 22-SEP-1999: 9908-0155139.
 FR 23-SEP-1999: 9908-0155139.
 FR 24-SEP-1999: 9908-0155139.
 FR 25-SEP-1999: 9908-0155139.
 FR 26-SEP-1999: 9908-0155139.
 FR 27-SEP-1999: 9908-0155139.
 FR 28-SEP-1999: 9908-0155139.
 FR 29-SEP-1999: 9908-0155139.
 FR 30-SEP-1999: 9908-0155139.
 FR 01-OCT-1999: 9908-0155139.
 FR 02-OCT-1999: 9908-0155139.
 FR 03-OCT-1999: 9908-0155139.
 FR 04-OCT-1999: 9908-0155139.
 FR 05-OCT-1999: 9908-0155139.
 FR 06-OCT-1999: 9908-0155139.
 FR 07-OCT-1999: 9908-0155139.
 FR 08-OCT-1999: 9908-0155139.
 FR 09-OCT-1999: 9908-0155139.
 FR 10-OCT-1999: 9908-0155139.
 FR 11-OCT-1999: 9908-0155139.
 FR 12-OCT-1999: 9908-0155139.
 FR 13-OCT-1999: 9908-0155139.
 FR 14-OCT-1999: 9908-0155139.
 FR 15-OCT-1999: 9908-0155139.
 FR 16-OCT-1999: 9908-0155139.
 FR 17-OCT-1999: 9908-0155139.
 FR 18-OCT-1999: 9908-0155139.
 FR 19-OCT-1999: 9908-0155139.
 FR 20-OCT-1999: 9908-0155139.
 FR 21-OCT-1999: 9908-0155139.
 FR 22-OCT-1999: 9908-0155139.
 FR 23-OCT-1999: 9908-0155139.
 FR 24-OCT-1999: 9908-0155139.
 FR 25-OCT-1999: 9908-0155139.
 FR 26-OCT-1999: 9908-0155139.
 FR 27-OCT-1999: 9908-0155139.
 FR 28-OCT-1999: 9908-0155139.
 FR 29-OCT-1999: 9908-0155139.
 FR 30-OCT-1999: 9908-0155139.

Query Match: 35.48; Score 29; DB 20; Length 9;
 Best Local Similarity: 55.68; Prod. No. 3.26+05;
 Matches: 5; Conservative: 2; Mismatches: 2; Indels: 0; Gaps: 0;

QY 2 ARDLTPKHVS 13
 ID 1 1 1 1 1 1 1 1 1

RESULT 14
 Y55614 standard; peptide; 9 AA.
 XX Y55614;
 AC Y55614;
 XX 17-JAN-2000 (first entry)
 DT HLA binding plu-1 peptide.
 DE Human cancer-associated polypeptide plu-1; ovarian cancer vaccine;
 EW breast cancer; immune response; cytotoxic T lymphocyte; imadag agent;
 KW therapeutic; plu-1 antigen; breast tumor; tumor antigen; epitope.
 XX Synthetic.
 OS Homo sapiens.
 XX W09949034-A1.
 PN 40 SEP-1999.
 XX 19-MAR-1999: 99W0-0800866.
 XX 20-MAR-1999: 99W0-0800867.
 XX (1999) INTERNATIONAL CARPES RES. TECHNOLOGY.
 XX Taylor papadimitriou J;
 XX WPI: 1999-591090/50.
 XX New nucleic acid encoding the cancer associated polypeptide plu-1, for
 diagnosis, treatment and prevention of cancer, especially of breast and
 ovary.
 Example 2: Fig 12; 173pp; English.
 XX The invention relates to a human cancer-associated polypeptide plu-1. The
 plu-1 polypeptide can be recombinantly expressed by standard recombinant
 methodology. Detection of the plu-1 nucleic acid or the polypeptide is
 used for the following: (i) diagnosis (including imaging) and prognosis
 of, and determination of susceptibility to, cancer, specifically ovarian
 or breast cancer; and (ii) treating cancer (by inducing an immune response
 against cancer cells, e.g. as a vaccine, or by antisense inhibition).
 CC Antisense derived from the polypeptide are used to generate activated
 CC cytotoxic T lymphocytes, or dendritic cells, for subsequent return to
 CC the patient for treatment of cancer. The polypeptide may also be used to
 CC identify inhibitors of plu-1 activity. Fragments of the polypeptide, and
 CC antibodies raised against plu-1, are useful as assay and imaging agents,
 CC also therapeutically (to induce an anti-idiotypic response or where
 CC conjugated to cytotoxic agents). The plu-1 antigen is expressed more
 CC commonly in breast tumors than some known tumor antigens. Sequences
 CC Y55620-625 represent predicted peptides from the plu-1 polypeptide which
 CC may bind to the human class I alleles B27, A2, A3 and A11.
 XX Sequence: 9 AA;
 SO

Query Match: 35.48; Score 29; DB 20; Length 9;
 Best Local Similarity: 55.68; Prod. No. 3.26+05;
 Matches: 5; Conservative: 2; Mismatches: 2; Indels: 0; Gaps: 0;

QY 3 ARDLTPKHVS 11
 ID 1 1 1 1 1 1 1 1 1

RESULT 15
 R81715
 ID R81715 standard; peptide; 15 AA.
 XX R81715;
 AC R81715;


```

XX 19 MAY 1995 (first entry)
XX Candida albicans enolase peptide #6.
XX
XX Candida albicans enolase; lambda-qt11; serum; antibody; candidiasis;
XX Thrush.
XX Synthesis.
XX W06269.2 AL.
XX 05 OCT 1995.
XX
XX 27 MAR 1995: 95W0 A009176.
XX
XX 25 MAR 1994: 94A0 0004742.
XX (BYU ) UNIV CORNELL UNIVERSITY.
XX Franklyn KM. Warrington JK.
XX WPI: 1995 051294/45.
XX
XX Protein or peptide reactive with Candida albicans induced antibodies
XX also DNA and vectors, used to treat, prevent or diagnose C.
XX albicans infections
XX
XX claim 1: page 26: 3/99: English.
XX
XX peptides 681710-25 are derived from the Candida albicans enolase protein
XX (681706). the peptides can be used to diagnose, in a serum sample,
XX antibodies induced by a Candida albicans infection and hence the
XX infection itself. the gene encoding the enolase was isolated from a
XX Candida albicans cDNA expression library in lambda-qt11 using human sera from a
XX patient with a clinical history of vaginitis. The resultant clone
XX contained a 470 bp open reading frame encoding residues 94-249 (K81709).
XX This clone was used to screen a genomic library and obtain the full
XX length cDNA.
XX
XX Genotype: 10 AA.

```

```

Query Match: 35.4%; Score 29; 108 bp; Length 15;
Best Local Similarity: 62.5%; Prod. No. 892
Matches: 1; Conservation: 2; Mismatches: 1; Indels: 0; Gaps: 0;
05 2 ADALP1K 9
06 1 11111
06 4 ADALP1K 11

```

Search completed: May 1, 2001, 11:26:33
 Job Time: 100 sec

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000


```

C:Accession: P07074
K:Tsukagawa, A.; Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Matsui, T.; Watanabe, Y.
Electrophoresis 21: 1853-1871, 2000
A:Title: Proteomic analysis of mouse brain: two dimensional electrophoresis profiles of
A:Reference number: P07072
A:Accession: P07074
A:Molecule type: protein
A:Residues: 1-9 <LSU>
A:Experimental source: strain 667g/467; Str. no.: brain, cortex
C:Keywords: Brain; cerebral cortex

Query Match: 29.9%; Score 24; DB 2; Length 9;
Best Local Similarity 80.0%; Pred. No. 26+04;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DLRTP R
II III
DB 4 DLRTP R

RESULT 12
A:Accession: A56871
C:Species: Sus scrofa domestica (domestic pig)
C:Title: 05-Jan-1996 #sequence_revision 21 04 1996 #text_change 09 May 1997
C:Accession: A56871
K:Stallins-Mann, M.L.; Trout, W.F.; Roberts, R.M.
Mol. Reprod. 48: 998-1005, 1993
A:Title: Porcine uterine retinol-binding proteins are identical gene products to the sea
A:Reference number: A56871; MUID:9324392
A:Accession: A56871
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-18 <STA>
A:Experimental source: endometrium
A:Note: Sequence extracted from NBR1 backbone (NBR1E-132137)
C:Superfamily: lipocalin; lipocalin homology

Query Match: 29.9%; Score 24; DB 2; Length 18;
Best Local Similarity 41.7%; Pred. No. 8.4e+02;
Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 RQKTPKQVSE 14
II III
DB 2 RQKVSSEKVP 13

RESULT 13
A:Accession: S52457
C:Species: Homo sapiens (man)
C:Title: 08-May-1995 #sequence_revision 21 04 1995 #text_change 05 May 1999
C:Accession: S52457
K:Antyokastitis, A.; Leversha, H.A.; Ferguson, Smith, M.; Moschonas, M.K.
Submitted to the EMBL data library, March 1993
A:Description: A cosmid clone mapped to human chromosome 11p15 detects a 1q11 restriction
A:Reference number: S52455
A:Accession: S52457
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-22 <AGR>
A:Cross-references: EMBL:X72881; NID:q667002; PUD:CAA51393.1; PID:q667004

Query Match: 29.9%; Score 24; DB 2; Length 22;
Best Local Similarity 57.1%; Pred. No. 1e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 10 HKVSEQR 16
II III
DB 10 HSLGEQR 16

```

```

RESULT 14
A:Accession: A30123
C:Species: Homo sapiens (man)
C:Title: 21-Feb-1990 #sequence_revision 21-Feb-1990 #text_change 16-Aug-1996
C:Accession: A30123
K:Seretic, P.D.; Trell, P.C.; Wright, J.; Pras, M.; Prangione, R.
J. Clin. Invest. 83: 836-843, 1989
A:Title: Systemic senile amyloidosis. Identification of a new protein in (trans)hyret
opathy.
A:Reference number: A30123; MUID:89155905
A:Accession: A30124
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-27 <GDR>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match: 29.9%; Score 24; DB 2; Length 27;
Best Local Similarity 62.5%; Pred. No. 1.3e+04;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 TPBEKQSE 14
II III
DB 5 TQHSVSE 12

RESULT 15
A:Accession: A34244
C:Species: Rattus norvegicus (Norway rat)
C:Title: 07-Sep-1993 #sequence_revision 07-Sep-1993 #text_change 28-Feb-1997
C:Accession: A34244
K:Magnum, M.A.; Shelton, K.D.
J. Biol. Chem. 264: 15936-15942, 1989
A:Title: An alternate promoter in the glucokinase gene is active in the pancreatic islet
A:Reference number: A34244; MUID:89480186
A:Accession: A34244
A:Molecule type: mRNA
A:Residues: 1-28 <MAG>
A:Cross-references: DB:M25807
A:Note: the authors translated the codon CTG for residue 2 as Lys
C:Superfamily: hexokinase; hexokinase homology
C:Keywords: ATP; glycolysis; phosphatase

Query Match: 29.9%; Score 24; DB 2; Length 28;
Best Local Similarity 42.9%; Pred. No. 1.4e+04;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 KADQATPKKQSE 14
II III
DB 5 RAEAEATKKKQEQ 18

```

Search completed: May 1, 2001, 11:26:22
Job time: 127 sec

Genware version 4.5
Copyright (c) 1994-2000 Compugen Ltd.

QM Protein Protein search, using sw model

Run on: May 1, 2001, 11:27:42, Search file: 23_94_records
(without alignments)
12,879,911,116, cell updates/sec

Title: us-09-341-829a-5_copy_164_179

Sequence: 82

1 KAKUUKTKRKVSSE24 16

Search label: BLSDM62

Gapop 10.0, Gapext 9.5

Searches: 185,757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 118244

Minimum hit seq length: 0

Maximum hit seq length: 60

Post processing: Minimum Match 0%

Maximum Match 100%

Listed first 4% summaries

Database: Issued Patent S.A.A.*

1 3 42.7 22 6 5178661 5
2 3 41.5 19 1 08-06-640-897 2
3 3 39.0 28 1 08-08-341-394 20
4 3 37.6 28 1 08-08-259-858 20
5 3 36.0 24 1 08-08-445-315 20
6 3 34.0 28 2 08-08-344-139 20
7 3 33.0 28 5 PCT-0895-06649-20
8 3 31.8 12 1 08-08-548-540 132
9 3 30.8 12 5 PCT-0896-09809-142
10 3 29.3 13 5 5178661 17
11 3 27.6 17 6 5219837 8
12 3 26.6 18 3 08-08-147-592A-17
13 3 25.4 20 2 08-08-411-859 12
14 3 24.4 21 3 08-08-746-111 7
15 3 23.4 21 3 08-09-282-351A-42
16 3 22.4 21 4 08-09-282-351A-37
17 3 21.4 21 4 08-09-282-351A-40
18 3 20.4 19 2 08-08-318-856A-64
19 3 19.4 12 2 08-08-179-557 34
20 3 18.4 12 4 08-08-981-601 6
21 3 17.4 16 2 08-08-179-557 37
22 3 16.4 15 4 08-09-981-601 12
23 3 15.4 20 1 08-08-399-696-79
24 3 14.4 20 2 08-08-447-420A-28
25 3 13.4 21 3 08-08-746-111 8
26 3 12.4 21 4 08-09-282-351A-44
27 3 11.4 21 4 08-09-282-351A-48

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | Id | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 45 | 42.7 | 22 | 6 | 5178661 5 |
| 2 | 34 | 41.5 | 19 | 1 | 08-06-640-897 2 |
| 3 | 32 | 39.0 | 28 | 1 | 08-08-341-394 20 |
| 4 | 32 | 37.6 | 28 | 1 | 08-08-259-858 20 |
| 5 | 32 | 36.0 | 24 | 1 | 08-08-445-315 20 |
| 6 | 32 | 34.0 | 28 | 2 | 08-08-344-139 20 |
| 7 | 32 | 33.0 | 28 | 5 | PCT-0895-06649-20 |
| 8 | 31 | 31.8 | 12 | 1 | 08-08-548-540 132 |
| 9 | 31 | 30.8 | 12 | 5 | PCT-0896-09809-142 |
| 10 | 31 | 29.3 | 13 | 5 | 5178661 17 |
| 11 | 30 | 27.6 | 17 | 6 | 5219837 8 |
| 12 | 30 | 26.6 | 18 | 3 | 08-08-147-592A-17 |
| 13 | 30 | 25.4 | 20 | 2 | 08-08-411-859 12 |
| 14 | 29 | 24.4 | 21 | 3 | 08-08-746-111 7 |
| 15 | 29 | 23.4 | 21 | 3 | 08-09-282-351A-42 |
| 16 | 29 | 22.4 | 21 | 4 | 08-09-282-351A-37 |
| 17 | 29 | 21.4 | 21 | 4 | 08-09-282-351A-40 |
| 18 | 28 | 20.4 | 19 | 2 | 08-08-318-856A-64 |
| 19 | 28 | 19.4 | 12 | 2 | 08-08-179-557 34 |
| 20 | 28 | 18.4 | 12 | 4 | 08-08-981-601 6 |
| 21 | 28 | 17.4 | 16 | 2 | 08-08-179-557 37 |
| 22 | 28 | 16.4 | 15 | 4 | 08-09-981-601 12 |
| 23 | 28 | 15.4 | 20 | 1 | 08-08-399-696-79 |
| 24 | 28 | 14.4 | 20 | 2 | 08-08-447-420A-28 |
| 25 | 28 | 13.4 | 21 | 3 | 08-08-746-111 8 |
| 26 | 28 | 12.4 | 21 | 4 | 08-09-282-351A-44 |
| 27 | 28 | 11.4 | 21 | 4 | 08-09-282-351A-48 |

28 34.1 28 1 08-07-729-348 56
29 34.1 28 1 08-07-967-421 56
30 34.1 28 1 08-08-240-141 56
31 34.1 28 1 08-08-548-540 56
32 34.1 28 5 PCT-0896-09809 56
33 32.9 9 2 08-08-418-856A 14
34 32.9 12 3 08-08-167-867 1
35 32.9 19 3 08-08-167-867 2
36 32.9 19 5 PCT-0895-06677 26
37 32.9 23 2 08-08-537-750B 242
38 32.9 23 3 08-08-871-355A 242
39 32.9 26 2 08-08-762-106 43
40 31.7 17 2 08-08-453-476 72
41 31.7 18 2 08-08-894-439 4
42 31.7 18 4 08-09-306-044 4
43 31.7 20 1 08-08-043-760 1
44 31.7 20 1 08-08-300-049 1
45 31.7 20 4 08-09-060-410 26

ALIGNMENTS

RESULT 1
5178661-5
Patent No. 5178661
APPLICANT: VERCARA, OLIVIERO; RUIZ, ANDRES; FERREIRA, ADELINO;
RUSSWELT, RUTH S; RUSSELSWEIG, VICTOR R.
TITLE OF INVENTION: CROSS-REACTIVE AND PROTECTIVE PEPTIDES
OF CIRCUMPORIZOITE PROTEINS
NUMBER OF SEQUENCES: 18
SUBJECT: ALLELATION DATA
ALLOCATION: 6-8-9608-1-1-241
FILING DATE: 12 JUN 1989
ALLOCATION DATA:
ALLOCATION NUMBER: 115,644
FILING DATE: 26 OCT 1987
APPLICATION NUMBER: 649,904
FILING DATE: 12 SEP 1984
SEQ ID NO: 5
LENGTH: 22
5178661 5

Query Match 42.7% Score 35; DB 6; Length 22;
Best Local Similarity 63.6%; Prod. No. 6,5;
Matches 7; Conservative 1; Mismatches 3; Gaps 0;

QY 5 LRTWKRVSEQ 15
ID 1 LRTWKRVSEQ 11

RESULT 2
08-06-640-897 2
Sequence 2, Application 05/086-0897
Patent No. 5659601

GENERAL INFORMATION:
APPLICANT: SAKEL, LAKATY
APPLICANT: SAKAMOTO, KENICHI
TITLE OF INVENTION: METHOD FOR DIAGNOSIS OF FOOT AND MOUTH
DISEASE, AND ANTIGENS CONTAINING THE PEPTIDE FOR DIAGNOSIS
TITLE OF INVENTION: OF FOOT AND MOUTH DISEASE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: SAKEL, SEIVAK, HOSUJARI, MATSE & NISHIDA
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk


```

RESULT 5
US 09 341 179 20
Sequence 20: Application US/087441 9
Patent No. 066912
GENERAL INFORMATION:
APPLICANT: Genentech, David V.
ATTORNEY/AGENT: Pottery, Mike
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/341 179
FILING DATE: 31 Oct 1996
CLASSIFICATION: 4.35
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250858
FILING DATE: 05/27/1994
ATTORNEY/AGENT INFORMATION:
NAME: Drosch, Glover R.
REGISTRATION NUMBER: 33,055
REFERENCE/BOOKLET NUMBER: 080701
TELEPHONE: 415/225 3216
TELEFAX: 415/952 9881
TELEX: 910/471-7168
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: Amino Acid
Topology: Linear
US 09 341 179 20

Query Match 99.08% Score 421 DB 1 Length 28
Best Local Similarity 42.98% Prod. No. 273
Matches 62 Conservation 120 2: Mismatches 6: Indels 0: Gaps 0:

27 4 MULTIPLEXSEQ 16
3 1 1 1 1 1
DB 13 KEEPTABLEVXOR 26

RESULT 7
US 09 341 179 20
Sequence 20: Application US/087441 9
Patent No. 066912
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb disk
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/087441 9
FILING DATE: 05/27/1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/250858
FILING DATE: 27 MAY 1994
ATTORNEY/AGENT INFORMATION:
NAME: Drosch, Glover R.
REGISTRATION NUMBER: 33,055
REFERENCE/BOOKLET NUMBER: 087441 9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225 3216
TELEFAX: 415/952 9881
TELEX: 910/471-7168
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: Amino Acid
Topology: Linear
US 09 341 179 20

Query Match 99.08% Score 421 DB 1 Length 28
Best Local Similarity 42.98% Prod. No. 273
Matches 62 Conservation 120 2: Mismatches 6: Indels 0: Gaps 0:

27 4 MULTIPLEXSEQ 16
3 1 1 1 1 1
DB 13 KEEPTABLEVXOR 26

RESULT 5
US 09 341 179 20
Sequence 20: Application US/087441 9
Patent No. 066912
GENERAL INFORMATION:
APPLICANT: Genentech, David V.
ATTORNEY/AGENT: Pottery, Mike
TITLE OF INVENTION: Tumor Necrosis Factor Receptor-Associated Factors
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA

```

TELEFAX: 415/952-9881
 TELEX: 910/571-7169
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 28 amino acids
 TYPE: amino acid
 YIELD: 1936000
 PC1-US96-06669-20

Query Match 49.0% Score 32; DB 5; Length 28;
 Best Local Similarity 42.9% Pred. No. 27;
 Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 8 RUDIPFHVSEQR 16
 Db 11111111

PC1-US96-540-540-142

Sequence 142, Application US/08/48549
 Patent No. 578731

GENERAL INFORMATION:

APPLICANT: Schatz, Peter J.

APPLICANT: Cull, Millard G.

APPLICANT: Miller, Jeff F.

APPLICANT: Stemmer, Willem P.C.

APPLICANT: Gates, Christian M.

TITLE OF INVENTION: Peptide Library and Screening Method

NUMBER OF SEQUENCES: 142

CORRESPONDENCE ADDRESS:

ADDRESSEE: William M. Smith

STREET: One Market Plaza, Stewart Tower, Suite 2000

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94105

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US 08/48549

FILING DATE: 26 OCT 1995

CLASSIFICATION: 446

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/290,641

FILING DATE: 15-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/963,321

FILING DATE: 15 OCT 1992

ADDRESS/AGENT INFORMATION:

NAME: Smith, William M.

REGISTRATION NUMBER: 30,223

REFERENCE/SEQ ID NO: 142981 03124005

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-426-2400

TELEFAX: 415-426-2422

INFORMATION FOR SEQ ID NO: 142:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

STRANDNESS: single

Topology: linear

MOLECULE TYPE: peptide

PC1-US96-540-540-142

Query Match 47.8% Score 31; DB 1; Length 12;
 Best Local Similarity 100.0% Pred. No. 17;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 HKVSEQ 15
 Db 4 HKVSEQ 9

RESULT 9

PC1-US96-09809-132

Sequence 132, Application PC/TUS96/09809

GENERAL INFORMATION:

APPLICANT: Schatz, Peter J.

APPLICANT: Cull, Millard G.

APPLICANT: Miller, Jeff F.

APPLICANT: Stemmer, Willem P.C.

APPLICANT: Gates, Christian M.

TITLE OF INVENTION: Peptide Library and Screening Method

NUMBER OF SEQUENCES: 142

CORRESPONDENCE ADDRESS:

ADDRESSEE: William M. Smith

STREET: One Market Plaza, Stewart Tower, Suite 2000

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94105

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US 08/290,641

FILING DATE: 15-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/963,321

FILING DATE: 15-OCT-1992

ADDRESS/AGENT INFORMATION:

NAME: Smith, William M.

REGISTRATION NUMBER: 30,223

REFERENCE/SEQ ID NO: 142981 03124005

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-426-2400

TELEFAX: 415-426-2422

INFORMATION FOR SEQ ID NO: 142:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

STRANDNESS: single

Topology: linear

MOLECULE TYPE: peptide

PC1-US96-09809-132

Query Match 47.8% Score 31; DB 5; Length 12;
 Best Local Similarity 100.0% Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 HKVSEQ 15
 Db 4 HKVSEQ 9

RESULT 10

PC1-US96-09809-132

Sequence 132, Application PC/TUS96/09809

GENERAL INFORMATION:

APPLICANT: Schatz, Peter J.

APPLICANT: Cull, Millard G.

APPLICANT: Miller, Jeff F.

APPLICANT: Stemmer, Willem P.C.

APPLICANT: Gates, Christian M.

TITLE OF INVENTION: Peptide Library and Screening Method

NUMBER OF SEQUENCES: 142

CORRESPONDENCE ADDRESS:

ADDRESSEE: William M. Smith

STREET: One Market Plaza, Stewart Tower, Suite 2000

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94105

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US 08/290,641

FILING DATE: 15-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/963,321

FILING DATE: 15-OCT-1992

ADDRESS/AGENT INFORMATION:

NAME: Smith, William M.

REGISTRATION NUMBER: 30,223

REFERENCE/SEQ ID NO: 142981 03124005

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-426-2400

TELEFAX: 415-426-2422

INFORMATION FOR SEQ ID NO: 142:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

STRANDNESS: single

Topology: linear

MOLECULE TYPE: peptide

PC1-US96-09809-132

Query Match 47.8% Score 31; DB 5; Length 12;
 Best Local Similarity 100.0% Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

TOP OF DOCUMENT: 115, 115, 115, 115

1 NUMBER OF SEQUENCES: 15
2 CURRENT APPLICATION DATA:
3 APPLICANT: BELL, GORDON, 05/05/1987, 241
4 FILING DATE: 22 JUN 1989
5 PRIOR APPLICATION DATA:
6 APPLICANT: BELL, GORDON, 11/5, 644
7 FILING DATE: 25 OCT 1987
8 APPLICATION NUMBER: 649,904
9 FILING DATE: 12 SEP 1984
10 SEQ ID NO: 17
11 LENGTH: 13
12 521064/ 17

Query Match: 67.6% Score 41 DB 6: Length 16
Best Local Similarity: 60.0% Prod. No. 187
Matches: 6: Conserved 1: Mismatches 4: Gaps 0:

27 6 RPKHVSSEQ 15
DB 1 111111111
2 EXPERRK 16

RESULT 11
521064/ 8
Patent No. 5210647
1 APPLICANT: BELL, GORDON, MARK L. WILLIAMS,
2 WILLIAM V.
3 TITLE OF INVENTION: METHOD OF STIMULATING MYELINATION
4 OF CELLS
5 NUMBER OF SEQUENCES: 9
6 CURRENT APPLICATION DATA:
7 APPLICANT: BELL, GORDON, 05/05/1987, 241, 775
8 FILING DATE: 21 JUN 1989
9 SEQ ID NO: 8
10 LENGTH: 17
11 521064/ 8

Query Match: 67.6% Score 40 DB 6: Length 17
Best Local Similarity: 60.6% Prod. No. 463
Matches: 6: Conserved 1: Mismatches 2: Gaps 0:

27 9 RPKHVSSEQ 15
DB 1 111111111
2 EXPERRK 16

RESULT 12
US 06 147 592A 17
1 Sequence 17: Application US/06147592A
2 Patent No. 6065433
3 GENERAL INFORMATION:
4 APPLICANT: BELL, GORDON, 1
5 APPLICANT: BELL, GORDON, 1
6 APPLICANT: Yasuda, Kazuki
7 TITLE OF INVENTION: Method of Receptor Genes,
8 TITLE OF INVENTION: Method of Receptor Genes,
9 NUMBER OF SEQUENCES: 43
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: Atsushi, Wada & Burke
12 STREET: P.O. Box 4433
13 CITY: Houston
14 STATE: Texas
15 COUNTRY: United States of America
16 ZIP: 77210
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Floppy disk
19 MEDIUM TYPE: IBM PC compatible
20 OPERATING SYSTEM: PC DOS/MS-DOS
21 SOFTWARE: Patent In Release #1.0, Version #1.25
22 CURRENT APPLICATION DATA:
23 APPLICANT: BELL, GORDON, 05/05/1987, 241, 775
24 FILING DATE: 21 JUN 1989
25 SEQ ID NO: 8
26 LENGTH: 17
27 521064/ 8

Query Match: 67.6% Score 40 DB 6: Length 16
Best Local Similarity: 60.0% Prod. No. 187
Matches: 6: Conserved 1: Mismatches 4: Gaps 0:

1 APPLICATION NUMBER: 05/05/147, 592A
2 FILING DATE: 05 NOV 1994
3 CLASSIFICATION: 435
4 PRIOR APPLICATION DATA:
5 APPLICATION NUMBER: 06/100, 694
6 FILING DATE: 30 JUL 1994
7 CLASSIFICATION: 435
8 APPLICATION NUMBER: 06/066, 296
9 FILING DATE: 20 MAY 1994
10 CLASSIFICATION: 435
11 ATTORNEY/AGENT INFORMATION:
12 NAME: Wilson, Mark H.
13 REGISTRATION NUMBER: 47,259
14 REFERENCE/DOCKET NUMBER: AKD-105
15 TELECOMMUNICATION INFORMATION:
16 TELEPHONE: (512) 418 4000
17 TELEFAX: (512) 474-7577
18 TELEX: N/A
19 INFORMATION FOR SEQ ID NO: 17:
20 SEQUENCE CHARACTERISTICS:
21 LENGTH: 16 amino acids
22 TYPE: amino acid
23 TOPOLOGY: linear
24 MOLECULE TYPE: protein
25 US-06-147-592A-17

Query Match: 67.6% Score 40 DB 6: Length 16
Best Local Similarity: 76.0% Prod. No. 483
Matches: 6: Conserved 1: Mismatches 2: Gaps 0:

27 1 KALDRTP 8
DB 1 111111111
2 KALDRTP 18

RESULT 13
US-08-411-859-12
1 Sequence 12: Application US/08411859
2 Patent No. 5985600
3 GENERAL INFORMATION:
4 APPLICANT: EVANS, CHRISTOPHER J.
5 APPLICANT: KEITH JR., DEANE E.
6 APPLICANT: EDWARDS, ROBERT H.
7 TITLE OF INVENTION: EXPRESSION CLONING OF A DELTA OPIOID
8 TITLE OF INVENTION: RECEPTOR RELATED EXPRESSION SYSTEMS, AND RELATED
9 TITLE OF INVENTION: PHARMACEUTICALS
10 NUMBER OF SEQUENCES: 15
11 CORRESPONDENCE ADDRESS:
12 ADDRESSEE: Morrison & Foster
13 STREET: 755 Page Mill Road
14 CITY: Palo Alto
15 STATE: California
16 COUNTRY: USA
17 ZIP: 94304 1018
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 MEDIUM TYPE: IBM PC compatible
21 OPERATING SYSTEM: PC DOS/MS-DOS
22 SOFTWARE: Patent In Release #1.0, Version #1.25
23 CURRENT APPLICATION DATA:
24 APPLICANT: Morrison & Foster, 08/08/411, 859
25 FILING DATE:
26 CLASSIFICATION: 424
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER: US 07/929, 200
29 FILING DATE: 13 AUG 1992
30 ATTORNEY/AGENT INFORMATION:
31 NAME: LITHGOW, TIMOTHY J.
32 REGISTRATION NUMBER: 46,856
33 REFERENCE/DOCKET NUMBER: 22300, 20526, 00
34 TELECOMMUNICATION INFORMATION:
35 TELEPHONE: 415-813-5600

1 TELEFAX: 415-494-0792
 2 TELEX: 706141
 3 INFORMATION FOR SEQ ID NO: 12:
 4 SEQUENCE CHARACTERISTICS:
 5 LENGTH: 20 amino acids
 6 TYPE: amino acid
 7 STRANDEDNESS: single
 8 TOPOLOGY: linear
 9 US-09-411-859-12

Query Match 46.6%; Score 30; DB 2; Length 20;
 Best Local Similarity 75.0%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KADLRTP 8
 11 11
 16 5 KALDRTP 12

RESULT 14
 US-09-745-111-7

1 Sequence 7, Application US/06746111
 2 Patent No. 4056778
 3 GENERAL INFORMATION:
 4 APPLICANT: Ginsburg, David
 5 APPLICANT: Cui, Jisong
 6 TITLE OF INVENTION: Compositions And Methods For Screening
 7 TITLE OF INVENTION: Compounds For Anticoagulant Activity
 8 NUMBER OF SEQUENCES 52
 9 CORRESPONDENCE ADDRESS:
 10 ADDRESSEE: Medline & Cartoll, LLP
 11 STREET: 220 Montgomery Street, Suite 2200
 12 CITY: San Francisco
 13 STATE: California
 14 COUNTRY: United States of America
 15 ZIP: 94104
 16 COMPUTER READABLE FORM:
 17 MEDIUM TYPE: Floppy disk
 18 COMPUTER: IBM PC compatible
 19 OPERATING SYSTEM: PC-DOS/MS-DOS
 20 SOFTWARE: Patent In Release #1.0, Version #1.40
 21 CURRENT APPLICATION DATA:
 22 APPLICATION NUMBER: 05/097746,111
 23 FILING DATE: 06-NOV-1996
 24 CLASSIFICATION:
 25 ATTORNEY/AGENT INFORMATION:
 26 NAME: Toroitai, Diane E.
 27 REGISTRATION NUMBER: 40,927
 28 REFERENCE/EXCERPT NUMBER: 00 02536
 29 INFORMATION INFORMATION:
 30 TELEPHONE: (415) 705-8410
 31 TELEFAX: (415) 497-8438
 32 INFORMATION FOR SEQ ID NO: 7:
 33 SEQUENCE CHARACTERISTICS:
 34 LENGTH: 21 amino acids
 35 TYPE: amino acid
 36 STRANDEDNESS: not relevant
 37 TOPOLOGY: not relevant
 38 MOLECULE TYPE: protein
 39 US-09-745-111-7

Query Match 45.4%; Score 29; DB 4; Length 21;
 Best Local Similarity 40.0%; Pred. No. 66;
 Matches 6; Conservative 2; Mismatches 7; Indels 7; Gaps 0;

QY 2 ARDLTKRKVSQR 16
 3 11
 16 5 AKTNPKKLRQR 21

RESULT 15

US-09-282-351A-42
 1 Sequence 42, Application 05/09282351A
 2 Patent No. 6194141

3 GENERAL INFORMATION:
 4 APPLICANT: Paul, Aulko V.
 5 APPLICANT: Wimmer, Eckard
 6 APPLICANT: Kiedor, Elizabeth
 7 TITLE OF INVENTION: INHIBITION OF PICOORNAVIRUS GENE REPLICATION BY
 8 TITLE OF INVENTION: INTERFERENCE WITH VPG NUCLEOTIDHYDRATION AND ELONGATION
 9 FILE REFERENCE: The Research Foundation of Stonybrook
 10 CURRENT APPLICATION NUMBER: 05/09282351A
 11 CURRENT FILING DATE: 1999-03-31
 12 NUMBER OF SEQ ID NOS: 49
 13 SOFTWARE: Patent In Ver. 2.1
 14 SEQ ID NO 32
 15 LENGTH: 21
 16 TYPE: PRT
 17 ORGANISM: Unknown organism
 18 FEATURES:
 19 OTHER INFORMATION: Description of Unknown organism: PICOORNAVIRUS
 20 US-09-282-351A-42

Query Match 45.4%; Score 29; DB 4; Length 21;
 Best Local Similarity 66.7%; Pred. No. 66;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 PDKRVSEQR 16
 11 11
 16 5 PDKRVSEQR 17

Search completed: May 1, 2001, 11:27:43
 Job time: 153 sec

PT 26-NOV-1992.
 XX
 PF 20-MAY-1992; 92W0-0804266.
 XX
 XX 21-MAY-1991; 910S-0704459.
 PF 12-JUL-1991; 910S-0724254.
 PF 27-JUN-1991; 915S-0734122.
 PF 28-AUG-1991; 910S-0751356.
 PF 20-SEP-1991; 915S-0762674.
 PF 14-NOV-1991; 910S-0791924.
 XX
 PF (GENE) GENENON PHARM INC.
 XX
 PF Hallbook F, Doney M, Liner CP, Peterson BB, Yanoopolous AB.
 XX
 PF WP1: 1992-415468/50.
 DE N-PSDB: 932248.
 XX
 XX Use of neurotrophin-4 for promoting growth and survival of nerve
 PF cells - useful in treating neurological, fertility and
 PF immunological disorders and in diagnosis
 XX
 PF Example 9; Page 74 - 135; 180pp; English.
 XX
 XX Example 9 describes the isolation and characterization of an NF 4
 CC human genomic clone. Three clones were found to hybridise to the
 CC NF-4 160 bp probe (see 829764 65) and the 45-mer probe given in
 CC Q2236. Corresp. to the amino acid sequence coded for by human per
 CC fragment clone 17B.
 XX
 XX Sequence 15 AA;
 SQ
 Query Match 32.7%; Score 35; DB 13; Length 15;
 Best Local Similarity 85.7%; Pred. No. 75;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 7 GAGGAGAG 14
 DB 1 11111
 1 GAGGAG 7
 RESULT 7
 Y54948
 ID Y54948 standard; peptide; 16 AA.
 AC Y54948;
 XX 15-FEB-2000 (first entry)
 XX
 DE Peptide ligand for fibrin polymerisation site.
 XX
 KW Thrombus imaging; fibrin polymerisation site; technetium 99m; Te 99m;
 KW fibrin/IIa receptor; cyclic peptide ligand.
 XX
 OS Synthetic.
 XX US596476 A.
 IN 19-067 1999.
 PF 07-JUN-1995; 950S-0484774.
 XX
 XX 11-JUL-1994; 940S-0274274.
 PF 21-MAY-1992; 92S-0906952.
 XX
 PF (GAT) GATHE INC.
 XX
 PF Dean RL, Lister-Jones J;
 XX
 DE WP1: 2000-021744/92
 XX
 PF A complex used for thrombus imaging comprises technetium 99m complexed

PT with a peptide ligand for GPIIb/IIIa receptor -
 XX
 PS disclosure; Column 8; 18pp; English.
 XX
 CC This sequence represents a peptide ligand for the fibrin polymerisation
 CC site. The invention relates to a complex (A) for thrombus imaging
 CC comprises technetium 99m (Tc 99m) complexed with a reagent comprising a
 CC peptide (P) with 4 to 100 amino acids sequence and a Te-99m binding
 CC moiety covalently bound to (P). (P) is selected from a linear peptide
 CC ligand for a GPIIb/IIIa receptor not comprising the amino acid sequence
 CC (arginine-glycine-aspartate), a peptide ligand for a polymerisation
 CC site of fibrin, and a cyclic peptide ligand for the GPIIb/IIIa receptor.
 CC The thrombus imaging reagents provided by the present invention can be
 CC used for visualising thrombi in a mammalian body when Tc-99m is labelled.
 XX
 XX Sequence 16 AA;
 SQ
 Query Match 42.7%; Score 45; DB 21; Length 16;
 Best Local Similarity 60.0%; Pred. No. 80;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 3 EXAGGAGGRC 12
 DB 5 54948 14
 3 54948 14
 RESULT 8
 W52103
 ID W52103 standard; peptide; 17 AA.
 AC W52103;
 XX 24-SEP-1998 (first entry)
 XX
 DE Targeting peptide #41 useful as component of thrombolytic agent.
 XX
 KW Thrombolytic agent; thrombolytic proteinase; blood clot; fibrin;
 KW thrombus; antithrombotic activity.
 XX
 OS Synthetic.
 XX W09824917 A1.
 PN 11-JUN-1998.
 PD 02-DEC-1997; 97W0-US21918.
 PF 02-DEC-1996; 960S-0754781.
 PR (DIAT-) DIATIDE INC.
 PA (CYSC-) UNIV SOUTHERN CALIFORNIA.
 XX
 PF Bush LR, Flores-Sanchez S E, Markland ES, Swenson S;
 XX
 DE WP1: 1998-433346/29.
 XX
 XX New thrombolytic agents - comprise thrombolytic proteinase
 PF covalently linked to targeting compound for binding to component of
 PF thrombus
 XX
 XX Claim 10; Page 64; 70pp; English.
 PS
 XX The invention relates to new thrombolytic agents which comprise a
 CC thrombolytic proteinase covalently linked to a targeting compound
 CC capable of specifically binding to a component of a thrombus. The
 CC thrombolytic agents can be used for eliminating thrombi in vivo in, e.g.,
 CC myocardial infarction, cerebral ischaemia, deep vein thrombosis or
 CC pulmonary embolism. A labelled form of the thrombolytic agent can also
 CC be used to image thrombi for diagnostic purposes. The thrombolytic
 CC agents are specifically targeted to thrombus sites in vivo and have
 CC minimal haemorrhagic side effects and side effects related to non-
 CC specific proteolysis. The present sequence represents a specifically

CV claimed targetting peptide.

XX Sequence 17 AA:

Query Match 42.7% Score 45; DB 19; Length 17;
Best Local Similarity 60.0%; Prod. No. 94;
Matches 6; Conservative 2; Mismatches 2; Gaps 0;

Q/ 4 EGAAGGCGG 12
: : : : :
D/ 5 dqrqdgrrg 14

RESULT 9

1D 960604 standard; peptide: 19 AA.

XX A:

XX 960604;

XX 24 JUN 1995 (first entry)

XX 0p 1H/111a receptor ligand used in scintigraphic imaging of thrombi.

XX Scintigraphic thrombus imaging; Specific binding;

XX Technetium-99m radiolabelled; 0p 11b/111a receptor ligand.

XX Synthetic;

XX Key Location/Qualifiers

FT Modified site 1 /note "S-acetamidomethyl Cys"

FT Modified site 4 /note "S-acetamidomethyl Cys"

XX Work done A.

XX 25 NOV 1993.

XX 21 MAY 1994; 9505-0504794.

XX 21 MAY 1994; 9205-0886742.

XX (DATA) DIAPEP INP

XX beam RL; Lister-James J;

XX WPI: 1993-09229/48.

XX Reagent for scintigraphic imaging of thrombi with 99m technetium

XX comprises synthetic peptide which binds to thrombus, covalently

XX coupled to metal binding group rapidly cleared from blood and

XX tissue.

XX claim 19; Page 42; 6pp; English.

XX The invention relates to reagents for scintigraphic imaging of a

XX thrombus in vivo, comprising (A) a specific binding compound capable of

XX binding to at least one component of a thrombus, covalently linked

XX to (B) a technetium-99m-binding moiety. Specific peptides

XX constituting the reagents are claimed as new. The present peptide is

XX one such peptide, in which the Cys(Acm)-Gly Cys(Acm) moiety is the

XX 99m Tc binding moiety and the residue constitutes the thrombus-

XX binding component.

XX Sequence 19 AA:

Q/ 4 EGAAGGCGG 12

D/ 5 dqrqdgrrg 14

1b : : : : :
8 dqrqdgrrg 17

RESULT 10

W050587

1D W50587 standard; peptide: 19 AA.

XX A:

XX W50587;

XX 16-JUL-1998 (first entry)

XX GPIIb/IIIa receptor ligand with technetium 99m binding moiety.

XX Technetium-99m label; thrombus imaging; GPIIb/IIIa receptor; ligand;

XX binding moiety.

XX Synthetic;

XX Key Location/Qualifiers

FT Modified site 1 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

FT Modified site 4 /note "Side chain sulphur of Cys blocked by

XX OS Ootus apella.
 XX
 XX
 FH Key Location/Qualifiers
 FI Modified site 27
 FI Zcode "amidated"
 XX
 XX
 FN W09402018-A.
 XX
 XX 03 FEB 1994.
 XX
 XX 27 JUL 1994; 9405-0507044.
 XX
 XX 27 JUL 1994; 9405-0519741.
 XX 23 MAR 1994; 9405-0049778.
 XX (MEDICAL RES FOUND) OREGON.
 XX
 FH Li K, Nardella SP, Spindel ER, Vijayaraghavan S;
 XX
 XX W09402018-A.
 XX N-PS06; Q54074.
 XX
 XX Now bombesin-like acrosome-related peptides used to promote the
 FI acrosome reaction to promote fertilisation or to develop
 FI antagonists to inhibit fertilisation.
 XX
 FI Claim 10; Page 46; 6pp; English.
 XX
 XX The peptide is a bombesin-like peptide which is capable of
 XX promoting the acrosome reaction to promote fertilisation.
 XX Bombesin antagonists can be used to inhibit fertilisation.
 XX See also R47609-20.
 XX
 XX Sequence 22 AA.
 XX

Query Match 32.7%; Score 45; ID 15; Length 27;
 Best Local Similarity 42.9%; Pred. No. 1.2e+02;
 Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

XX 1. A0402018-VNVM 16
 XX 03 JUL 1994
 XX 04 SEP 1994 22

XX 24
 XX 04 JAN 1994 (first entry)
 XX
 XX Human B cell stimulator 2 (RSP2) antagonist.
 XX
 XX Immunodeficiency; myeloma; chronic articular rheumatism;
 XX endotoxin shock.
 XX
 XX Bone capons.
 XX
 XX J094020060-A.
 XX
 XX 24 JUL 1990.
 XX
 XX 17 JAN 1989; 89JP-0007044.
 XX
 XX 17 JAN 1989; 89JP-0007044.
 XX
 XX (CHS) CHS-AI PHARMACEUTICAL KK.
 XX
 XX W09402018-A.
 XX

PI Polypeptide with antagonism against human B cell stimulation 2
 PI comprises human B cell stimulator with deficiency of 2 or more
 PI aminoacid(s) at N and/or C-terminal.
 XX
 XX Claim 12; Page 969; 6pp; Japanese.
 XX
 XX RSP2 antagonists are useful in diagnosis and treatment of immune
 XX deficiencies e.g. myeloma, chronic articular rheumatism and endotoxin
 XX shock. Antagonists comprises RSP2 peptide (prot. AAs 20-40) with
 XX substitutions of two or more C- and/or N terminal AAs
 XX
 XX Sequence 25 AA;
 XX

Query Match 32.7%; Score 45; ID 15; Length 27;
 Best Local Similarity 46.7%; Pred. No. 1.2e+02;
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

XX 2. P02A020303CVAVN 16
 XX 13 JUL 1994
 XX 6 pmaackkddtqsdu 20

RES011 15
 R47609
 ID R47609 standard; peptide; 27 AA.
 XX
 XX R47609;
 XX
 XX 26-JUL-1994 (first entry)
 XX
 XX Bombesin-like peptide.
 XX
 XX Acrosome reaction; fertilisation.
 XX
 XX Ootus apella.
 XX
 XX Key Location/Qualifiers
 FI Modified site 27
 FI Zcode "amidated"
 XX
 XX W09402018-A.
 XX
 XX 04 FEB 1994.
 XX
 XX 27-JUL-1994; 9405-0507044
 XX
 XX 27-JUL-1994; 9205-0919741.
 XX 24-MAR-1994; 9405-0049778.
 XX
 XX (MEDICAL RES FOUND) OREGON
 XX
 FH Li K, Nardella SP, Spindel ER, Vijayaraghavan S;
 XX
 XX W09402018-A.
 XX N-PS06; Q54074.
 XX
 XX Now bombesin-like acrosome-related peptides used to promote the
 FI acrosome reaction to promote fertilisation or to develop
 FI antagonists to inhibit fertilisation
 XX
 FI Claim 10; Page 46; 6pp; English.
 XX
 XX The peptide is a bombesin-like peptide which is capable of
 XX promoting the acrosome reaction to promote fertilisation.
 XX Bombesin antagonists can be used to inhibit fertilisation.
 XX See also R47610-20.
 XX
 XX Sequence 27 AA;
 XX

Query Match 32.7%; Score 45; ID 15; Length 27;
 Best Local Similarity 42.9%; Pred. No. 1.2e+02;

Matches 6: Conservative 4: Mismatches 5: Indels 0: Gaps 0:

QY 5 ACHGGRGVAFVVM 18

1 1 1 1 1 1 1 1

DE 14 dsqgac1qtpqhlm 27

Search completed: May 1, 2001, 11:25:12
Job time: 10.2 sec

NAME: CLIFF, JAMES A
 REGISTRATION NUMBER: 27,075
 REFERENCE: 100 INCOMPATIBLE
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 846 6400
 TELEFAX: (703) 846 2787
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 24 amino acids
 TYPE: amino acid
 STRAIGHTNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHEetical: No
 US 08-441-008-2

Query Match: 41.1% Score 44; DB 4; Length 24;
 Best Local Similarity 58.4% Pred. No. 3.2;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 PPGAGAGGGRG 12
 1 1 1 1 1 1
 IR 5 PPGKRGGRG 16

RESULT 4
 US 08-440 594-8
 Sequence No. Application US/08440594-8
 Patent No. 5702506
 GENERAL INFORMATION:
 APPLICANT: CAREY, DARRELL H.
 TITLE OF INVENTION: SYNTHETIC PEPTIDE NEUTROTROPHIC CELL GROWTH-FACTIVE AGENTS
 FILE REFERENCE: CHN:002
 CURRENT APPLICATION NUMBER: 08-440 594-8
 CURRENT FILING DATE: 1994-10-28
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO: 8
 LENGTH: 24
 TYPE: PEI
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 US 08-440 594-8

Query Match: 41.1% Score 44; DB 4; Length 24;
 Best Local Similarity 58.4% Pred. No. 3.2;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 PPGAGAGGGRG 12
 1 1 1 1 1 1
 IR 5 PPGKRGGRG 16

RESULT 4
 US 08-441-947-4
 Sequence No. Application US/08451947
 Patent No. 5702506
 GENERAL INFORMATION:
 APPLICANT: GENENTECH, INC.
 TITLE OF INVENTION: NOVEL NEUTROTROPHIC FACTOR
 NUMBER OF SEQUENCES: 100
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 460 Point San Bruno Blvd
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA

ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC DOS/MS DOS
 SOFTWARE: Path (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 08/08451,947
 FILING DATE:
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/426419
 FILING DATE: 19-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/040013
 FILING DATE: 22 MAR 1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/48442
 FILING DATE: 31 JAN
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/487707
 FILING DATE: 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Torchia, Timothy E.
 REGISTRATION NUMBER: 36,700
 REFERENCE/AGENT NUMBER: 664F2-102
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415/225-8674
 TELEFAX: 415/262-9881
 INDEX: 910271 7153
 INFORMATION FOR SEQ ID NO: 40:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 80 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US 08-441-947-40

Query Match: 40.2% Score 43; DB 1; Length 80;
 Best Local Similarity 80.0% Pred. No. 5.8;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GAQGGGRGV 14
 1 1 1 1 1 1
 IR 12 GAQGGGRGV 21

RESULT 5
 US 08-451-947-4
 Sequence No. Application US/08451947
 Patent No. 5702506
 GENERAL INFORMATION:
 APPLICANT: GENENTECH, INC.
 TITLE OF INVENTION: NOVEL NEUTROTROPHIC FACTOR
 NUMBER OF SEQUENCES: 100
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.
 STREET: 460 Point San Bruno Blvd
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94080
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC DOS/MS DOS
 SOFTWARE: Path (Genentech)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 08/08451,947
 FILING DATE:
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:

1 APPLICATION NUMBER: 06/24,419
 2 FILING DATE: 19 APR 1995
 3 PRIOR APPLICATION DATA:
 4 APPLICATION NUMBER: 06/24,419
 5 FILING DATE: 22 MAR 1995
 6 PRIOR APPLICATION DATA:
 7 APPLICATION NUMBER: 07/64,482
 8 FILING DATE: 31 JAN 1995
 9 PRIOR APPLICATION DATA:
 10 APPLICATION NUMBER: 07/68,770
 11 FILING DATE: 1991
 12 ATTORNEY/AGENT INFORMATION:
 13 NAME: Torchia, PhD., Timothy E.
 14 REGISTRATION NUMBER: 46,700
 15 TELECOMMUNICATION INFORMATION:
 16 TELEPHONE: 415/225-9861
 17 TELEFAX: 415/225-9861
 18 TELEX: 910/71-7168
 19 INFORMATION FOR SEQ ID NO: 4:
 20 SEQUENCE CHARACTERISTICS:
 21 LENGTH: 40 amino acids
 22 TYPE: amino acid
 23 TOPOLOGY: Linear
 24 US 08 421 826A 43

Query Match 40.2% Score 43 DB 1 Length 40
 Best Local Similarity 80.0% Pred. No. 5,8; Mismatches 2; Gaps 0;
 Matches 6; Conserved 0

Q7 4 GAGDGGPGV 13
 11 11111
 14 11 GAGGGGPGV 20

RESULT 5
 US 08 421 826A 40
 : Sequence 43, Application US/08424826A
 : Patent No. 5830858
 : GENERAL INFORMATION:
 : APPLICANT: Rosenthal, Arnon
 : TITLE OF INVENTION: RAVEL NIBRO-PROTEIN FACTOR
 : NUMBER OF SEQUENCES: 98
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Genentech, Inc.
 : STREET: 460 Point San Bruno Blvd
 : CITY: South San Francisco
 : STATE: California
 : COUNTRY: USA
 : ZIP: 94080
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 5.5 inch, 1.44 Mb floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC DOS/MS DOS
 : SOFTWARE: WINDAT (Genentech)
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: 06/24,419
 : FILING DATE: 19 APR 1995
 : CLASSIFICATION: 514
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 07/64,482
 : FILING DATE: 31 JAN 1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Torchia, PhD., Timothy E.
 : REGISTRATION NUMBER: 46,700
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 415/225-9861
 : TELEFAX: 415/225-9861
 : TELEX: 910/71-7168
 : INFORMATION FOR SEQ ID NO: 4:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 40 amino acids
 : TYPE: Amino Acid
 : TOPOLOGY: Linear
 : US-08-424-826A-43

1 TELECOMMUNICATION INFORMATION:
 2 TELEPHONE: 415/225-9861
 3 TELEFAX: 415/225-9861
 4 TELEX: 910/71-7168
 5 INFORMATION FOR SEQ ID NO: 40:
 6 SEQUENCE CHARACTERISTICS:
 7 LENGTH: 30 amino acids
 8 TYPE: Amino Acid
 9 TOPOLOGY: Linear
 10 US-08-424-826A-40

Query Match 40.2% Score 43 DB 2 Length 40
 Best Local Similarity 80.0% Pred. No. 5,8; Mismatches 2; Gaps 0;
 Matches 6; Conserved 0

Q7 4 GAGDGGPGV 13
 11 11111
 14 12 GAGGGGPGV 21

RESULT 7
 US-08-424-826A 43
 : Sequence 43, Application US/08424826A
 : Patent No. 5830858
 : GENERAL INFORMATION:
 : APPLICANT: Rosenthal, Arnon
 : TITLE OF INVENTION: RAVEL NIBRO-PROTEIN FACTOR
 : NUMBER OF SEQUENCES: 98
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Genentech, Inc.
 : STREET: 460 Point San Bruno Blvd
 : CITY: South San Francisco
 : STATE: California
 : COUNTRY: USA
 : ZIP: 94080
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 5.5 inch, 1.44 Mb floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC DOS/MS DOS
 : SOFTWARE: WINDAT (Genentech)
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: 06/24,419
 : FILING DATE: 19 APR 1995
 : CLASSIFICATION: 514
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 07/64,482
 : FILING DATE: 31 JAN 1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Torchia, PhD., Timothy E.
 : REGISTRATION NUMBER: 46,700
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 415/225-9861
 : TELEFAX: 415/225-9861
 : TELEX: 910/71-7168
 : INFORMATION FOR SEQ ID NO: 43:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 40 amino acids
 : TYPE: Amino Acid
 : TOPOLOGY: Linear
 : US-08-424-826A-43

Query Match 40.2% Score 43 DB 2 Length 40
 Best Local Similarity 80.0% Pred. No. 5,8; Mismatches 2; Gaps 0;
 Matches 6; Conserved 0

```

QY 4 GAGGCGGCV 13
LB 11 GAGGCGGCV 20

RESULT 8
US-08-928-694-40
? Sequence 40, Application US/08/28694
? Patent No. 6037420
? GENERAL INFORMATION:
? APPLICANT: ROSENTHAL, ARNON
? TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
? NUMBER OF SEQUENCES: 100
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Genentech, Inc.
? STREET: 1 DNA Way
? CITY: South San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 94080
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: WinPatIn (Genentech)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/928,694
? FILING DATE: 12-Sep-1997
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/451947
? FILING DATE: 26-MAY-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/426419
? FILING DATE: 19-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/030013
? FILING DATE: 22-MAR-1993
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/648482
? FILING DATE: 31-JAN
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/587707
? FILING DATE: 1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Torchia, PhD., Timothy E.
? REGISTRATION NUMBER: 36,700
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 650/225-8674
? TELEFAX: 650/225-8674
? INFORMATION FOR SEQ ID NO: 40:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 40 amino acids
? TYPE: Amino Acid
? TOPOLOGY: Linear
US-08-928-694-40

```

```

Query Match 40.2% Score 43; DB 3; Length 30;
Best Local Similarity 80.0%; Pred. No. 5,8;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GAGGCGGCV 13
LB 12 GAGGCGGCV 21

RESULT 9
US-08-928-694-43
? Sequence 43, Application US/08/28694
? Patent No. 6037420

```

```

? GENERAL INFORMATION:
? APPLICANT: ROSENTHAL, ARNON
? TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
? NUMBER OF SEQUENCES: 100
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Genentech, Inc.
? STREET: 1 DNA Way
? CITY: South San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 94080
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: WinPatIn (Genentech)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/928,694
? FILING DATE: 12-Sep-1997
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/451947
? FILING DATE: 26-MAY-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/426419
? FILING DATE: 19-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/030013
? FILING DATE: 22-MAR-1993
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/648482
? FILING DATE: 31-JAN
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/587707
? FILING DATE: 1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Torchia, PhD., Timothy E.
? REGISTRATION NUMBER: 36,700
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 650/225-8674
? TELEFAX: 650/225-8674
? INFORMATION FOR SEQ ID NO: 43:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 40 amino acids
? TYPE: Amino Acid
? TOPOLOGY: Linear
US-08-928-694-43

```

```

Query Match 40.2% Score 43; DB 3; Length 30;
Best Local Similarity 80.0%; Pred. No. 5,8;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GAGGCGGCV 13
LB 11 GAGGCGGCV 20

```

```

RESULT 10
PCT-US91-06950-40
? Sequence 40, Application PCT/US9106950
? GENERAL INFORMATION:
? APPLICANT: ROSENTHAL, ARNON
? TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
? NUMBER OF SEQUENCES: 100
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Genentech, Inc.
? STREET: 460 Point San Bruno Blvd
? CITY: South San Francisco
? STATE: California
? COUNTRY: USA

```

```

1  ZIP: 94080
2  COMPUTER READABLE FORM:
3  MEDIUM TYPE: 5.25 inch, 660 Kb floppy disk
4  OPERATING SYSTEM: IBM PC compatible
5  SOFTWARE: pat lib (Gensoft)
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: 01/587797
8  FILING DATE: 19910924
9  CLASSIFICATION: 436
10  APPLICATION DATA:
11  APPLICATION NUMBER: 07/648482
12  APPLICATION DATA:
13  APPLICATION NUMBER: 01/587707
14  NAME: Bousley, Max D.
15  REGISTRATION NUMBER: 27,043
16  REFERENCE NUMBER: 66641
17  TELECOMMUNICATION INFORMATION:
18  TELEPHONE: 415/266-1994
19  TELEFAX: 415/266-9861
20  TELE: 415/266-7166
21  INFORMATION FOR SEQ ID NO: 40:
22  SEQUENCE CHARACTERISTICS:
23  LENGTH: 40 amino acids
24  TYPE: AMINO ACID
25  TOPOLOGY: linear
26  US91-06950-4

```

```

Query Match: 40.2% Score 43 DB 5; Length 40;
Best Local Similarity: 80.0% Prod. No. 5.8;
Matches 6; Conservat Ipe 0; Mismatches 2; Indels 0; Gaps 0;

```

```

27 4 GAGGAGRGV 14
28 1111111
29 12 GAGGAGRGV 21

```

```

RESULT 11
1  US91-06950-43
2  Sequence 43: Application patent 436
3  GENERAL INFORMATION:
4  APPLICANT: GENENTECH, INC.
5  APPLICANT: GENENTECH, INC.
6  TITLE OF INVENTION: N-VEL NIBB-4-PHIL-PAC-6
7  NUMBER OF SEQUENCES: 169
8  CORRESPONDENCE ADDRESS:
9  ADDRESSEE: Genentech, Inc.
10  STREET: 460 Point San Bruno Blvd
11  CITY: South San Francisco
12  STATE: California
13  COUNTRY: USA
14  ZIP: 94080
15  COMPUTER READABLE FORM:
16  MEDIUM TYPE: 5.25 inch, 660 Kb floppy disk
17  OPERATING SYSTEM: IBM PC compatible
18  SOFTWARE: pat lib (Gensoft)
19  CURRENT APPLICATION DATA:
20  APPLICATION NUMBER: 01/587797
21  FILING DATE: 19910924
22  CLASSIFICATION: 436
23  APPLICATION DATA:
24  APPLICATION NUMBER: 07/648482
25  APPLICATION DATA:
26  APPLICATION NUMBER: 01/587797
27  NAME: Bousley, Max D.
28  REGISTRATION NUMBER: 27,043
29  REFERENCE NUMBER: 66641
30  TELECOMMUNICATION INFORMATION:
31  TELEPHONE: 415/266-1994

```

```

1  TELEFAX: 415/266-9861
2  TELE: 415/266-7166
3  INFORMATION FOR SEQ ID NO: 43:
4  SEQUENCE CHARACTERISTICS:
5  LENGTH: 40 amino acids
6  TYPE: AMINO ACID
7  TOPOLOGY: linear
8  US91-06950-44

```

```

Query Match: 40.2% Score 43 DB 5; Length 40;
Best Local Similarity: 80.0% Prod. No. 5.8;
Matches 6; Conservat Ipe 0; Mismatches 2; Indels 0; Gaps 0;

```

```

29 4 GAGGAGRGV 14
30 1111111
31 12 GAGGAGRGV 20

```

```

RESULT 12
1  US-08-981-088-4
2  Sequence 4: Application 08-981088
3  Patent No. 6146924
4  GENERAL INFORMATION:
5  APPLICANT: HAR SHAVIT, RACHEL
6  TITLE OF INVENTION: ARTI METASTATIC ANT-ARTI ANTICORP
7  TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
8  NUMBER OF SEQUENCES: 4
9  CORRESPONDENCE ADDRESS:
10  ADDRESSEE: OLIFF & BERTRIDGE, PLLC
11  STREET: P.O. BOX 19928
12  CITY: ALEXANDRIA
13  STATE: VA
14  COUNTRY: USA
15  ZIP: 22320
16  COMPUTER READABLE FORM:
17  MEDIUM TYPE: floppy disk
18  OPERATING SYSTEM: IBM PC compatible
19  SOFTWARE: Patent In Release #1.0, Version #1.0
20  CURRENT APPLICATION DATA:
21  ATTORNEY/AGENT INFORMATION:
22  FILING DATE: 27 JAN-1998
23  CLASSIFICATION: 514
24  ATTORNEY/AGENT INFORMATION:
25  NAME: OLIFF, JAMES A
26  REGISTRATION NUMBER: 27,076
27  REFERENCE NUMBER: JAN 40455
28  TELECOMMUNICATION INFORMATION:
29  TELEPHONE: (703) 846-6400
30  TELEFAX: (703) 846-2787
31  INFORMATION FOR SEQ ID NO: 4:
32  SEQUENCE CHARACTERISTICS:
33  LENGTH: 14 amino acids
34  TYPE: amino acid
35  STRANDEDNESS: single
36  TOPOLOGY: linear
37  MOLECULE TYPE: peptide
38  HYDROTHERICAL: No
39  US-08-981-088-4

```

```

Query Match: 45.5% Score 48 DB 4; Length 14;
Best Local Similarity: 60.0% Prod. No. 14;
Matches 6; Conservat Ipe 1; Mismatches 5; Indels 0; Gaps 0;

```

```

29 1 PPGAGRGVAC 10
30 1111111
31 5 PPGAGRGVAC 14

```

```

RESULT 13
1  US-08-482-880-11

```

Sequence 11, Application US/09-482,480
Patent No. 5746122
GENERAL INFORMATION:
APPLICANT: Dean, Richard T.
ATTORNEY/AGENT INFORMATION:
TITLE OF INVENTION: Technetium 99m Labeled Peptides for
TITLE OF INVENTION: Thrombus Imaging
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allogretti, Ltd.
STREET: 10 South Wacker Drive Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPILED: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 95/09,482,480
FILING DATE: 07 JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 5746122, Kevin E
REGISTRATION NUMBER: 35,403
REFERENCE/CKET NUMBER: 92,216-M
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910 221-5317
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-482-480-11

Query Match: 42.7% Score 35; DB 1; Length 16;
Best Local Similarity: 60.0%; Pred. No. 40;
Matches: 6; Conservative: 2; Mismatches: 2; Indels: 0; Gaps: 0;
QY 3 DGGGDDGG 12
DB 5 DGGGDDGG 14

RESULT 14
US-09-274-274-11
Sequence 11, Application US/09-482,480
Patent No. 5746122
GENERAL INFORMATION:
APPLICANT: Dean, Richard T.
ATTORNEY/AGENT INFORMATION:
TITLE OF INVENTION: Technetium 99m Labeled Peptides for
TITLE OF INVENTION: Thrombus Imaging
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allogretti & Witte, Ltd.
STREET: 10 South Wacker Drive Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPILED: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

Sequence 11, Application US/09-482,480
Patent No. 5746122
GENERAL INFORMATION:
APPLICANT: Dean, Richard T.
ATTORNEY/AGENT INFORMATION:
TITLE OF INVENTION: Technetium 99m Labeled Peptides for
TITLE OF INVENTION: Thrombus Imaging
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allogretti, Ltd.
STREET: 10 South Wacker Drive Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPILED: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 95/09,482,480
FILING DATE: 07 JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 5746122, Kevin E
REGISTRATION NUMBER: 35,403
REFERENCE/CKET NUMBER: 92,216-M
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910 221-5317
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-274-274-11

Query Match: 42.7% Score 35; DB 2; Length 16;
Best Local Similarity: 60.0%; Pred. No. 40;
Matches: 6; Conservative: 2; Mismatches: 2; Indels: 0; Gaps: 0;
QY 3 DGGGDDGG 12
DB 5 DGGGDDGG 14

RESULT 15
US-09-475-041-11
Sequence 11, Application US/04475041
Patent No. 5879658
GENERAL INFORMATION:
APPLICANT: Dean, Richard T.
ATTORNEY/AGENT INFORMATION:
TITLE OF INVENTION: Technetium 99m Labeled Peptides for
TITLE OF INVENTION: Thrombus Imaging
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allogretti, Ltd.
STREET: 10 South Wacker Drive Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPILED: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 95/09,475,041
FILING DATE: 07 JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 5879658, Kevin E
REGISTRATION NUMBER: 35,403
REFERENCE/CKET NUMBER: 92,216-M
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312 715 1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-475-041-11

Query Match: 42.74; Score: 0; DB: 2; Length: 16;
 Best Local Similarity: 60.00; Pred. No: 40;
 Matches: 0; Conserved: 2; Mismatches: 2; Gaps: 0;

00: 5 0000000000000000
 10: 5 0000000000000000

Search completed: May 1, 2001, 11:27:32
 14.0000000000000000


```

Query Match: 36.6%; Score 30; DB 1; Length 22;
Best Local Similarity 42.9%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 0; Gaps 0;

QY * KRLRTPKRVKVSQR 16
II II I I I I
LI 1 KRLRTPKRVKVSQR 14

RESULT
CXM1 CONNE
ID CXM1 CONNE STANDARD; PRT; 22 AA;
AC P05492;
DT 01-NOV-1988 (Rel. 09; Created)
DI 01-NOV-1988 (Rel. 09; Last sequence update)
DE 01-APR-1990 (Rel. 15; Last annotation update)
DE MD-CONTOXIN-3116;
OS Conus geographus (Geography cone)
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoida; Conidae; Conus;
CX NCBI_FaxID 6491;
FN (1);
PP SEQUENCE;
RX MEDLINE 95261416; PubMed 2410412;
RA Gray W.R., Gray W.R., Morozowski E.J.,
RA Yoshikami D., Morozowski E.J.,
RA "Conus geographus toxins that discriminate between compound and
RA muscle sodium channels.";
RA J. Biol. Chem. 260:9280-9288(1985);
RN 121;
RP SEQUENCE;
RX MEDLINE 95261416; PubMed 2410412;
RA Sato S., Nakamura H., Ohizumi Y., Kobayashi J., Hirata Y.,
RA "The amino acid sequences of hemolysins hydroxyproline containing
RA myotoxins from the marine snail Conus geographus venom.";
RA FEBS Lett. 155:277-280(1984);
RN 131;
RP DISULFIDE BONDS;
RX MEDLINE 96249596; PubMed 2328142;
RA Hirata Y., Sato S., Nakamura H., Ohizumi Y.,
RA "Disulfide pairings in geographotoxin I, a peptide neurotoxin from
RA Conus geographus.";
RA FEBS Lett. 264:29-32(1990);
RN 141;
RP REVIEW;
RX MEDLINE 96249596; PubMed 2328142;
RA Gray W.R., Gray W.R., Morozowski E.J.,
RA "Peptide toxins from venomous Conus snails.";
RA Annu. Rev. Biochem. 57:665-700(1988);
RN 151;
RP STRUCTURE BY NMR;
RX MEDLINE 91122275; PubMed 1991506;
RA At E.H., Becker S., Gordon R.D., Roeterjans H.,
RA "Solution structure of mu-conotoxin GI1A analysed by 2D NMR and
RA dynamic nuclear polarization.";
RA FEBS Lett. 278:160-166(1991);
RN 161;
RP STRUCTURE BY NMR;
RX MEDLINE 91299744; PubMed 2069951;
RA Lancelini J.-M., Kohda D., Tate S.-I., Yamawaki Y., Abo T., Satoh M.,
RA Inaaki F.,
RA "Tertiary structure of conotoxin GI1A in aqueous solution.";
RA Biochemistry 30:6908-6916(1991);
CX * FUNCTION, MD-CONTOXINS ACTION, MUSCLE MEMBRANES, VERY BLACK
CX * MUSCLE CONTRACTION BY BINDING TO THE VOLTAGE ACTIVATED SODIUM
CX CHANNELS;
DR PIR; A01786; MXXNI;
DR PIR; A23579; A23579;
DR PDB; 1TCG; 31-JAN-94;
DR PDB; 1TCJ; 31-JAN-94;
DR PDB; 1TCJ; 31-JAN-94;
DR PDB; 1TCJ; 31-JAN-94;
KW Sodium channel inhibitor; Hydroxylation; Amidation; Venom;
KW 4D-structure;
FT DISULFID 4 15
FT DISULFID 4 20
FT DISULFID 10 21
FT MOD_RES 6 6
FT MOD_RES 7 7
FT MOD_RES 17 17
FT MOD_RES 22 22
FT HELIX 13 16
FT TURN 19 21
SQ SEQUENCE 22 AA; 2568 MW; 150492DAV2A98140 QRS64;

Query Match: 35.4%; Score 29; DB 1; Length 22;
Best Local Similarity 42.9%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY * KRLRTPKRVKVSQR 16
II II I I I I
LI 1 KRLRTPKRVKVSQR 14

RESULT
CXM1 CONNE
ID CXM1 CONNE STANDARD; PRT; 22 AA;
AC P05492;
DT 01-NOV-1988 (Rel. 01; Created)
DI 01-NOV-1988 (Rel. 01; Last sequence update)
DE 01-FEB-1995 (Rel. 31; Last annotation update)
DE MD-CONTOXIN-3116 (MY-TOXIN 1) (GE-GRAPHOTOXIN 1) (CTX 1);
OS Conus geographus (Geography cone)
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoida; Conidae; Conus;

```

DB 1 KAKUPTPKK 23

RESULT 4
 RL18 HALB
 ID RL18 HALB STANDARD: PRT: 22 AA.
 AC 105560;
 DT 01-OCT-1996 (Rel. 34, Created)
 DI 01-OCT-1996 (Rel. 34, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 50S RIBOSOMAL PROTEIN L18P (HME118) (FRAGMENT).
 GN RPL18P.
 OS Halobacterium mediterraneum (Halobacteriales; Halobacteriaceae; Haloterrax.
 AC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Haloterrax.
 CX NCBI TaxID: 2241;
 RN 11;
 RP SEQUENCE.
 RC STRAIN DSM 1411;
 RX MEDLINE 94229075; PubMed 8174557;
 RA McDonnell J., Wittmann-Liebold B.,
 RI "Comparative analysis of the protein components from 5S rRNA-protein
 RI complexes of halophilic archaebacteria.";
 RL Eur. J. Biochem. 221:779-785(1994).
 CY 1. SIMILARITY: RELATIONS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.
 DE PIR: R33084; R33084.
 KW Ribosomal protein.
 FT NON_TER 22 22
 FT SEQUENCE 22 AA: 2775 MW: 244519.07498746 Da; 22

Query Match 30.5% Score 25; DB 1; Length 22;
 Best Local Similarity 36.4% Prod. No. 2.2e+02;
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 1 KAKUPTPKK 11
 DB 12 KREVTTHQ 22

RESULT 5
 RL18 HALB
 ID RL18 HALB STANDARD: PRT: 23 AA.
 AC 105570;
 DT 01-OCT-1996 (Rel. 34, Created)
 DI 01-OCT-1996 (Rel. 34, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 50S RIBOSOMAL PROTEIN L18P (HME118) (FRAGMENT).
 GN RPL18P.
 OS Halobacterium mediterraneum (Halobacteriales; Halobacteriaceae; Haloterrax.
 AC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Haloterrax.
 CX NCBI TaxID: 2241;
 RN 11;
 RP SEQUENCE.
 RC STRAIN DSM 1411;
 RX MEDLINE 94229075; PubMed 8174557;
 RA McDonnell J., Wittmann-Liebold B.,
 RI "Comparative analysis of the protein components from 5S rRNA-protein
 RI complexes of halophilic archaebacteria.";
 RL Eur. J. Biochem. 221:779-785(1994).
 CY 1. SIMILARITY: RELATIONS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.
 DE PIR: R33084; R33084.
 KW Ribosomal protein.
 FT NON_TER 23 23
 FT SEQUENCE 23 AA: 2924 MW: 244519.07498746 Da; 23

Query Match 30.5% Score 25; DB 1; Length 23;
 Best Local Similarity 36.4% Prod. No. 2.2e+02;
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 1 KAKUPTPKK 11
 DB 12 KREVTTHQ 22

RESULT 6
 RL18 HALB
 ID RL18 HALB STANDARD: PRT: 26 AA.
 AC 105560;
 DT 01-OCT-1996 (Rel. 34, Created)
 DI 01-OCT-1996 (Rel. 34, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 50S RIBOSOMAL PROTEIN L18P (HME118) (FRAGMENT).
 GN RPL18P.
 OS Halobacterium mediterraneum (Halobacteriales; Halobacteriaceae; Haloterrax.
 AC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Haloterrax.
 CX NCBI TaxID: 2241;
 RN 11;
 RP SEQUENCE.
 RC STRAIN DSM 670;
 RX MEDLINE 94229075; PubMed 8174557;
 RA McDonnell J., Wittmann-Liebold B.,
 RI "Comparative analysis of the protein components from 5S rRNA-protein
 RI complexes of halophilic archaebacteria.";
 RL Eur. J. Biochem. 221:779-785(1994).
 CY 1. SIMILARITY: RELATIONS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.
 DE PIR: R33084; R33084.
 KW Ribosomal protein.
 FT NON_TER 26 26
 FT SEQUENCE 26 AA: 3269 MW: 156593.8746136074 Da; 26

Query Match 30.5% Score 25; DB 1; Length 26;
 Best Local Similarity 36.4% Prod. No. 2.2e+02;
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KAKUPTPKK 11
 DB 12 KREVTTHQ 22

RESULT 7
 RL18 HALB
 ID RL18 HALB STANDARD: PRT: 30 AA.
 AC 105570;
 DT 01-NOV-1988 (Rel. 09, Created)
 DI 01-NOV-1988 (Rel. 09, Last sequence update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 50S RIBOSOMAL PROTEIN L18P (HME118) (FRAGMENT).
 GN RPL18P.
 OS Halobacterium mediterraneum (Halobacteriales; Halobacteriaceae; Haloterrax.
 AC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Haloterrax.
 CX NCBI TaxID: 2240;
 RN 11;
 RP SEQUENCE.
 RC MEDLINE 79045279; PubMed 152199;
 RX Smith R., Matheson A.L., Yatchew R., Willick G., Ragan R.N.,
 RA "The 5S rRNA protein complex from an extreme halophile,
 RI Halobacterium cutitubum: purification and characterization.";
 RL Eur. J. Biochem. 89:501-509(1978).
 CY 1. SIMILARITY: RELATIONS TO THE L18P FAMILY OF RIBOSOMAL PROTEINS.
 DE PIR: R07277; R07277.
 KW Ribosomal protein.
 FT NON_TER 30 30
 FT SEQUENCE 30 AA: 3624 MW: 385007.9315661074 Da; 30

Query Match 30.5% Score 25; DB 1; Length 30;
 Best Local Similarity 36.4% Prod. No. 3e+02;
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KAKUPTPKK 11
 DB 12 KREVTTHQ 22

```

RESULT  8
TRYD_FR_A1
ID TRYP_FR_A1 STANDARD: PRI: 21 AA.
AC P12229;
DT 01-FEB-1994 (Rel. 29, Created)
DT 01-FEB-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1995 (Rel. 42, Last annotation update)
DE TRYPSIN PRECURSOR (EC 3.4.21.4) (FRAGMENT).
OS Protophytus actinophilus (Marbled lungfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Dipnoi; Lepidosireniformes; Protopteriidae; Protopterus.
CX NCBI_TaxID:7886;
RN [1]
RP SEQUENCE.
RC TISSUE:Pancreas;
RA Herzigson M.A., Iye K.W., Beck S.K., Reutath B., Walsh K.A.,
RA "Comparison of the amino terminal sequences of bovine, dogfish, and
RA lungfish trypsins.",
RL FEBS Lett. 14:222-224(1971).
CC 1: CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC 2: SUBCELLULAR LOCATION: EXTRACELLULAR.
CC 3: SIMILARITY: BELONGS TO THE TRYPSIN FAMILY 51. AKA: KN-WN AS THE
CC TRYPSIN FAMILY.
LF PIR: A2719; A2719.
DR BSSP: P07288; LPFA.
DR MER-05: S01151;
DR InterPro: IPR001264;
DR PROSITE: PS00144; TRYPSIN_HIS: PARTIAL.
DR PROSITE: PS00145; TRYPSIN_SER: PARTIAL.
KW Hydrolase; Serine protease; Zymogen.
FI Peptid 1 7 ACTIVATION PEPTIDE.
FI CHAIN 2 21 257526.
FI RS-LER 23 21
SQ SEQUENCE 21 AA: 2454 MW: 398270.107230 CR64.

Query Match 28.8%; Score 24; DB 1; Length 21;
Best Local Similarity 80.0%; Pred. No. 4.6e-02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 8 PRKRV 12
DB 15 PRKRV 19

RESULT  9
PSAL_PEA
ID PSAL_PEA STANDARD: PRI: 13 AA.
AC P17229;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-FEB-1998 (Rel. 37, Last annotation update)
DE PROTEIN SYSTEM 1 REACTION CENTRE SUBUNIT IX (PSI IX) (FRAGMENT).
GN PSAL.
OS Pisum sativum (Garden pea).
OC Chloropast.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta.
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 1;
OC Fabales; Fabaceae; Papilionoideae; Pisum.
CX RIV_TaxID: 6088;
RN [1]
RP SEQUENCE.
RC MEDLINE 90242987; PubMed 2185954;
RA Pechen M., Hirata A., Miyama T., Inoue Y.
RA "Polypeptide composition of higher plant photosystem I complex.
RA Identification of psal1, psal2 and psal3 gene products.",
RL FEBS Lett. 264:274-278(1990).
CC 1: FUNCTION: MAY HELP IN THE ORGANIZATION OF THE PSAL AND PSAP
CC SUBUNIT.
CC 2: SIMILARITY: BELONGS TO THE PSAL FAMILY.
DR PIR: S09743; S09743.
KW Chloroplast; Photosystem I; Photosynthesis; Transmembrane.

```

```

ET TRANSMEM 6 13 POTENTIAL.
ET NON_TER 13 13
SQ SEQUENCE 13 AA: 1516 MW: 952645.011103841 CR64;

Query Match 26.8%; Score 22; DB 1; Length 13;
Best Local Similarity 21.4%; Pred. No. 9.4e-02;
Matches 3; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 3 ROLKT 7
DB 2 ROLKT 6

RESULT 10
HORC_HORSP
ID HORC_HORSP STANDARD: PRI: 28 AA.
AC P02864;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE c HORDEIN (FRAGMENT).
OC Hordeum spontaneum (Barley).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae; Triticeae;
OC Hordeum.
CX NCBI_TaxID:77009;
RN [1]
RP SEQUENCE.
RA Shewry P.R., Autran J.-C., Nimmo C.C., Lew E.J., Kasarda D.D.,
RA "N terminal amino acid sequence homology of storage protein components
RA from barley and a diploid wheat.",
RL Nature. 284:520-522(1980).
CC 1: FUNCTION: SWEETEN PEEK SHEP; STARCH PROTEIN.
CC 2: TISSUE SPECIFICITY: DEVELOPING; EMBOSPERM.
DR PIR: A03355; A03355.
KW Seed storage protein; Multiques family.
ET NON_TER 28 28
SQ SEQUENCE 28 AA: 4333 MW: 808420.244775AA CR64;

```

```

Query Match 26.8%; Score 22; DB 1; Length 28;
Best Local Similarity 21.4%; Pred. No. 9.4e-02;
Matches 3; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 2 AEDLTPKHKVSQ 15
DB 7 SRILOSPQSYLQQ 20

RESULT 11
V39_SPV4
ID V39_SPV4 STANDARD: PRI: 28 AA.
AC P11241;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-JUL-1989 (Rel. 11, Last annotation update)
DE GENE 9 PECTEN.
GN 9.
OS Spiroplasma virus 4 (SPV4).
OC Viruses; ssDNA viruses; Microviridae; Spiromicrovirinae.
CX NCBI_TaxID 10855;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE 99022909; PubMed 2922658;
RA Renaudin J., Pascari M.C., Rove J.M.
RA "Spiroplasma virus 4, nucleotide sequence of the viral RNA,
RA regulatory signals, and proposed genome organization.",
RL J. Bacteriol. 169:4950-4961(1987).
CC 1: THIS SVS3 PECTEN IS VIRULIFEROUS. IT IS PRODUCED THROUGH A COLLABORATION
CC BETWEEN THE GENE INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION
CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS

```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed, change by and for commentariat.org. All other uses require a license agreement (see <http://www.commentariat.org/licenses/18b34b.ch>).

EMOL; M/T9BB; ; NO1_ANNOTATE_CDS,
FIR; B2002; ; 194BSW,
SHEATH; 2B AA; 3775 MW; 9916090409F1110 CROG4;

UNIVERSITY MICROFILMS
 300 N. ZEEB RD.
 ANN ARBOR, MI 48106-1500
 TEL: 734/769-0900
 FAX: 734/769-0901
 WWW: WWW.UMI.EDU

| Matches | 4; counter 94179 | 1; Missed calls | 2; Models | 0; caps | 0; |
|---------|------------------|-----------------|-----------|---------|----|
| 27 | 6; PLEKHA7 | 12 | | | |

71. AHHZAAH 011 1977

| | |
|----------------|----------------------|
| HPLC Buffer A | |
| Urea | HPLC Buffer A |
| N ₂ | 100% 16; |
| | STANDARD; PR; 29 AA. |

01 FEB 1994 (001, 17, Last sequence update)
01 NOV 1995 (001, 32, Last annotation update)
01 DEC 1995 (001, 34, 24, 50) (HEMORRHAGIC METALLOPROTEINASE J)

Bohrerops jararaca (Jatavaca).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Lepidosteiformes; Synbranchii; Siluriformes; Siluriformes; Colubridae.

| | |
|---|--|
| | Violet color; λ_{max} at 604m μ ; Red dyes. |
| X | NH ₄ acetate; 97.44; |
| X | NH ₄ acetate; 97.44; |
| N | [] |
| N | [] |
| N | Silver nitrate. |

[illegible]

¹⁴For illustration and some characteristics of a zinc metalloprotease from the genus of *Rothornis* (Gardner), *ibid.*, *Expt. 2*, 747-755 (1989).

14 ALA + 150-15 OF INSULIN B CHAIN AND PRO-1 OF ARGENTINEN IN
1 CATALYTIC ACTIVITY; CLEAVAGE OF 4 CYS-1 HIS-5, 9 SER-1 HIS-10 AND
1 ACTS IN HEMOPHAGET, ENZYME RELEASE FROM SINGLE-CELL CULTURE

ANAL. Calcd for $C_{10}H_{10}O_2$: C, 80.0%; H, 8.0%. Found: C, 79.8%; H, 8.2%. IR (KBr): 1715 (C=O), 1640 (C=C), 1600 (C=C), 1580 (C=C), 1540 (C=C), 1520 (C=C), 1500 (C=C), 1480 (C=C), 1460 (C=C), 1440 (C=C), 1420 (C=C), 1400 (C=C), 1380 (C=C), 1360 (C=C), 1340 (C=C), 1320 (C=C), 1300 (C=C), 1280 (C=C), 1260 (C=C), 1240 (C=C), 1220 (C=C), 1200 (C=C), 1180 (C=C), 1160 (C=C), 1140 (C=C), 1120 (C=C), 1100 (C=C), 1080 (C=C), 1060 (C=C), 1040 (C=C), 1020 (C=C), 1000 (C=C), 980 (C=C), 960 (C=C), 940 (C=C), 920 (C=C), 900 (C=C), 880 (C=C), 860 (C=C), 840 (C=C), 820 (C=C), 800 (C=C), 780 (C=C), 760 (C=C), 740 (C=C), 720 (C=C), 700 (C=C), 680 (C=C), 660 (C=C), 640 (C=C), 620 (C=C), 600 (C=C), 580 (C=C), 560 (C=C), 540 (C=C), 520 (C=C), 500 (C=C), 480 (C=C), 460 (C=C), 440 (C=C), 420 (C=C), 400 (C=C), 380 (C=C), 360 (C=C), 340 (C=C), 320 (C=C), 300 (C=C), 280 (C=C), 260 (C=C), 240 (C=C), 220 (C=C), 200 (C=C), 180 (C=C), 160 (C=C), 140 (C=C), 120 (C=C), 100 (C=C), 80 (C=C), 60 (C=C), 40 (C=C), 20 (C=C), 0 (C=C).

| | |
|-----|---------------------------------------|
| 109 | HSS10: $[1.54, 1.79]$; $2\Delta[10]$ |
| 109 | MEPS10: $[1.2, 1.49]$ |
| 109 | 109-10: $[1.90, 0.05]$ 402 |
| 109 | 109-05: $[0.05, 0.05]$ 402 |
| 109 | 109-05: $[0.05, 0.05]$ 402 |

| | |
|----|---|
| KW | Hydro-Québec; Metalloprotection; Zink; Vermin. |
| FI | NIN 1ER ²⁹ |
| NO | SEPIKENT ²⁹ AA; 34.3 MW; C26FHC5A13E34E C9564; |

| | | | |
|-----------------------|--------|--------------|-------------------|
| anion Match | | | |
| Root Local Similarity | 26.9% | Serie: 22; | IdB 1; Length 27; |
| | 60.0%; | Prod. No. 9, | Bos02; |

[illegible]
$$\begin{array}{llll} \text{I} \text{A} \text{I} & \text{H} \text{V} \text{I} \text{Z}^4 & & \\ \vdots & \text{I} \text{A} \text{I} & \text{H} \text{V} \text{I}^4 & \\ \text{A}^4 & \text{P} \text{I}^4 & \text{C}^4 \text{I} \text{O} & \\ & \text{S} \text{I} \text{A} \text{N} \text{I} \text{A} \text{P} \text{I}^4 & \text{P} \text{R} \text{I} & \text{Z}^4 \text{A} \text{A}^4 \end{array}$$

FT N_NLTER 14 14
 SQ SEQUENCE 14 AA: 1756 MW: 65583660 PERC25B CRC64;

Query Match: 25.6%; Score 21; DB 1; Length 14;
 Best Local Similarity: 42.9%; Pred. No. 6,9e-02;
 Matches: 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

CV 4 LRTPKH 10
 LB 7 EYHAPKH 14

RESULT 15

HEXADENOV
 11 HEXADENOV STANDARD; PRT; 18 AA.
 AC P45985;
 DT 01-JUN-1994 (Ref. 29, Created)
 DT 01-JUN-1994 (Ref. 29, Last sequence update)
 DT 01-NOV-1997 (Ref. 35, Last annotation update)
 DE HEXON PROTEIN (LATE PROTEIN 2) (FRAGMENT).
 IN F11.
 OS Canine adenovirus type 1 (strain Utrecht).
 NC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 CC NBI_Lax10 46364;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE:90043182; PubMed 1413543;
 RA Cal P.; Weber J.M.;
 RI "Nucleotide and deduced amino acid sequence of the canine adenovirus
 type 1 proteinase."
 RL Virus Genes 6:407-412(1992).
 SC 1- JUNE 1994; THIS PROTEIN IS ONE OF THE STRUCTURAL PROTEINS IN THE
 1- VIRAL CAPSID AND IS SYNTHESIZED DURING LATE INFECTION.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. Please contact: helpdesk@ebi.ac.uk
 CC or send an email to license@ebi.ac.uk.
 CC
 LR EMBL: M72715; AAA2528.1; .
 LR EMBL: A48550; A48550.
 LR HSSP: P04277; IDHX.
 KW Coat protein; Hexon protein; Late protein.
 FT N_NLTER 1 1
 SQ SEQUENCE 18 AA: 1949 MW: 19924607A120962 CRC64;

Query Match: 25.6%; Score 21; DB 1; Length 18;
 Best Local Similarity: 100.0%; Pred. No. 8,9e-02;
 Matches: 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CV 5 LRTPKH 8
 LB 7 LRTPKH 10

Search completed: May 1, 2001, 11:25:52
 Job time: 102 sec

01-MAY-2000 (TrEMBLrel. 14, last annotation update)
 DE ANTON EX-HAMMER AIR 4 PROTEIN (FRAGMENT).
 OS Gallus gallus (Chicken).
 CA Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archaeosauria; Aves; Neognathae; Caprimulgidae; Phasianidae; Phasianinae.
 GN Gallus.
 CX NPTL_TaxID 9041;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 95288119; PubMed 7408967;
 RA Cox K.H., Cox J.V.;
 RI "Viruses: Retroviral proteins possess divergent NH(2)-terminal
 cytoplasmic domains".
 RL Am. J. Physiol. 268:0-0(0).
 DR EMBL: S76800; AAC14248.1; F.
 FI N_N_TER 1
 SQ SEQUENCE 28 AA; 2412 MW; 94989E22040B236 CRG64;

Query Match 66.7%; Score 24; DB 14; Length 28;
 Best Local Similarity 83.4%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Q7 1 REGAGR 6
 DB 1 REGAGR 6
 RESULT 4
 Q9QVQ9
 ID Q9QVQ9 PRELIMINARY: PRT: 20 AA.
 AC Q9QVQ9
 DT 01-MAY-2000 (TrEMBLrel. 14, Created)
 DI 01-MAY-2000 (TrEMBLrel. 14, last sequence update)
 DE ALPHA CLASS GLUTATHIONE S-TRANSFERASE SUBUNIT 2 (P24472)
 NE (FRAGMENT).
 OS Rattus sp.
 CA Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Rodentia; Muridae; Murinae; Eutamias.
 GN Rattus.
 CX NPTL_TaxID 10118;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE 96045981; PubMed 7485987;
 RA Ewing P., Bracker L., Tuller J.;
 RI "Electrospray ionization-mass spectrometry as a tool for
 characterization of glutathione S-transferase isozymes".
 RL Anal. Biochem. 229:404-412(1995).
 DR BSSP: P24472; IGDK.
 SQ SEQUENCE 20 AA; 2412 MW; 94989E22040B236 CRG64.

Query Match 63.9%; Score 24; DB 14; Length 20;
 Best Local Similarity 66.7%; Pred. No. 2.2e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Q7 2 REGAGR 7
 DB 10 Q9GRM 15
 RESULT 4
 Q9QVQ9
 ID Q9QVQ9 PRELIMINARY: PRT: 21 AA.
 AC Q9QVQ9
 DT 01-MAY-2000 (TrEMBLrel. 14, Created)
 DI 01-MAY-2000 (TrEMBLrel. 14, last sequence update)
 DE ALPHA CLASS GLUTATHIONE S-TRANSFERASE SUBUNIT 2 (P24472)
 NE (FRAGMENT).
 OS Rattus sp.
 CA Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Rodentia; Muridae; Murinae; Eutamias.
 GN Rattus.
 CX NPTL_TaxID 10118;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE 96045981; PubMed 7485987;
 RA Ewing P., Bracker L., Tuller J.;
 RI "Electrospray ionization-mass spectrometry as a tool for
 characterization of glutathione S-transferase isozymes".
 RL Anal. Biochem. 229:404-412(1995).
 DR BSSP: P24472; IGDK.
 SQ SEQUENCE 20 AA; 2412 MW; 94989E22040B236 CRG64.

CX NPTL_TaxID 10118;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE 96045981; PubMed 7485987;
 RA Ewing P., Bracker L., Tuller J.;
 RI "Electrospray ionization-mass spectrometry as a tool for
 characterization of glutathione S-transferase isozymes".
 RL Anal. Biochem. 229:404-412(1995).
 DR BSSP: P24472; IGDK.
 SQ SEQUENCE 21 AA; 2475 MW; AFA6BA961E24972C CRG64;

Query Match 63.9%; Score 23; DB 11; Length 21;
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Q7 2 REGAGR 7
 DB 11 Q9GRM 16
 RESULT 5
 Q9QVQ9
 ID Q9QVQ9 PRELIMINARY: PRT: 27 AA.
 AC Q9QVQ9
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DI 01-JAN-1998 (TrEMBLrel. 05, last sequence update)
 DE 01-NOV-1998 (TrEMBLrel. 08, last annotation update)
 NE ENVELOPE GLYCOPROTEIN (FRAGMENT).
 OS Human immunodeficiency virus type 1.
 CA Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 GN HIV-1.
 CX NPTL_TaxID 11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 96881399; PubMed 9443234;
 RA Nelson J.A.E., Fiscus S.A., Swainston R.;
 RI "Isolation, characterization and immunodetection of the human immunodeficiency virus type 1 V
 region characterised by using a laboratory kit assay".
 RL J. Virol. 71:8750-8758(1997).
 DR EMBL: AF001804; AAC58485.1;
 KW Envelope protein.
 FT N_N_TER 1
 FT N_C_TER 27
 SQ SEQUENCE 27 AA; 2944 MW; 1274B084E89FA60 CRG64;

Query Match 63.9%; Score 24; DB 14; Length 27;
 Best Local Similarity 83.4%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Q7 1 REGAGR 6
 DB 9 REGAGR 14
 RESULT 6
 Q9QVQ9
 ID Q9QVQ9 PRELIMINARY: PRT: 28 AA.
 AC Q9QVQ9
 DT 01-MAY-2000 (TrEMBLrel. 14, Created)
 DI 01-MAY-2000 (TrEMBLrel. 14, last sequence update)
 DE GLUTATHIONE S-TRANSFERASE SUBUNIT YX, GST SUBUNIT YX
 OS Rattus sp.
 CA Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Rodentia; Sciurognathii; Muridae; Murinae; Rattus.
 GN Rattus.
 CX NPTL_TaxID 10118;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE 92233942; PubMed 1567476;
 RA Igarashi T., Tsuchiya T., Shikata Y., Sakami F., Igarashi T., Horie T.

10 11:11
11 11:11:18

```

RESULT 15
Q0XBRZ
AC Q0XBRZ PRELIMINARY: PRT: 24 AA.
D1 01 NOV 1999 (REMBLRL: 12, Created)
D1 01 NOV 1999 (REMBLRL: 12, Last sequence update)
D1 01 NOV 1999 (REMBLRL: 12, Last annotation update)
DE FRAMBLATION ELONGATION FACTOR IS (EF TS) (FRAMBLMENT).
GN ISF.
CN Bacillus cereus.
CU Bacteria, Firmicutes, Bacillus/Clostridium group.
CA Bacillus/Staphylococcus group: Bacillus.
AX N601 0410 1996.
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN A100 1996.
RX MEDLINE 04241048; PubMed 10217496.
RA Kistner A., Boud L., Lindhaek T., Rishard A.L., Kistner A.L.
RI "Sequence annotation is not conserved between Bacillus cereus and
R1 Bacillus subtilis."
R1 Microbiology 14:1621-631 (1999).
R1 EMBL: A10104; 07040592.1;
KW Elongation factor.
C1 N N 199 24
C2 SEQUENCE 24 AA; 2671 MW; F1A990F0B550ABD CRO64;
Query Match.
Best Local Similarity: 67.18; Score 21; ID: 2; Length 24;
Matches: 4; Conservation: 1; Mismatches: 2; Gaps: 0;
Q7 1 PPTAGRM 7
C1 110.1
C2 14 K1GAMM 20

```

Search completed: May 1, 2001, 11:27:04
 Search time: 144 sec

RESULT 3
 PQ0143
 polydactylonase (EC 3.2.1.15) p26 - evening primrose (fragment)
 C:Species: crotonera organoculus (evening primrose)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 06-Dec-1996
 C:Accession: PQ0143
 R:Brown, S.M.; Crouch, M.L.
 Plant Cell 2: 253-274, 1990
 A:Title: Characterization of a gene family abundantly expressed in crotonera organoculus
 A:Reference number: J00992; MUID:93005668
 A:Accession: PQ0143
 A:Molecule type: mRNA
 A:Residues: 1-22 <BB>
 A:Experimental source: pollen
 C:Comment: This protein is specifically translated in the pollen.
 C:Comment: This protein functions by dephosphorylating protein in the cell walls of the plant
 growing tube.
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation

 Query Match 27.1% Score 29; Db 2; Length 22;
 Best Local Similarity 93.3% Pred. No. 2.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 4 GAGGAG 9
 L L L L L
 Lb 15 GAGGAG 20

 RESULT 4
 A55244
 tubulin alpha-2 chain - Paramesidium tetraurelia (fragment)
 C:Species: Paramesidium tetraurelia
 C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 07-Dec-1999
 C:Accession: A55244
 R:Schacter, V.; Levilliers, N.; Schmitter, J.M.; Le Goff, J.P.; Rossier, J.; Adolphe,
 Science 266: 1688-1691, 1994
 A:Title: Polyglycylation of tubulin: a posttranslational modification in axonemal microtubules
 A:Reference number: A55244; MUID:95084156
 A:Accession: A55244
 A:Molecule type: protein
 A:Status: preliminary
 C:Comment: This protein is a small, water-soluble apolipoprotein, is thought to increase
 is a residual pathway for flight-related lipid transport.
 C:Keywords: hemolymph; lipid transport
 C:Genetic code: SGG5
 C:Superfamily: tubulin

 Query Match 25.2% Score 27; Db 2; Length 28;
 Best Local Similarity 62.5% Pred. No. 7.4e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

 QY 5 AAGAGGEC 12
 L L L L L L
 Lb 10 AAGAGGEC 17

 RESULT 5
 A55244
 tubulin alpha-1 chain - Paramesidium tetraurelia (fragment)
 C:Species: Paramesidium tetraurelia
 C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 07-Dec-1999
 C:Accession: A55244
 R:Schacter, V.; Levilliers, N.; Schmitter, J.M.; Le Goff, J.P.; Rossier, J.; Adolphe,
 Science 266: 1688-1691, 1994
 A:Title: Polyglycylation of tubulin: a posttranslational modification in axonemal microtubules
 A:Reference number: A55244; MUID:95084156
 A:Accession: A55244
 A:Molecule type: protein
 A:Status: preliminary
 C:Comment: This protein is a small, water-soluble apolipoprotein, is thought to increase
 is a residual pathway for flight-related lipid transport.
 C:Keywords: hemolymph; lipid transport
 C:Genetic code: SGG5
 C:Superfamily: tubulin

A:Residues: 1-28 <BB>
 R:Blomback, B.; Blomback, M.; Grundahl, N.J.
 Acta Chem. Scand. 19: 1789-1791, 1965
 A:Title: Studies on fibrinopeptides from mammals.
 A:Reference number: A03118
 A:Accession: A03118
 A:Molecule type: protein
 A:Residues: 1-16 <BB>
 R:Schubert, J.; A.J.; Colman, R.W.; Laki, K.; Gladner, J.A.
 Biochem. Biophys. Res. Commun. 14: 555-558, 1964
 A:Reference number: A37511; MUID:66020594
 A:Accession: A37511
 A:Molecule type: protein
 A:Residues: 1-16, EKG 16-16 <BB>
 C:Superfamily: fibrinogen alpha chain; fibrinogen disulfide ring homology
 C:Keywords: blood coagulation; liver; phosphoprotein; plasma
 F1-16/Protein; fibrinopeptide A #status: experimental
 F1-16/Protein; fibrinopeptide A (partial) #status: experimental

 Query Match 25.2% Score 27; Db 2; Length 28;
 Best Local Similarity 62.5% Pred. No. 7.4e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

 QY 5 AAGAGGEC 12
 L L L L L L
 Lb 10 AAGAGGEC 17

 RESULT 6
 A60295
 apolipoprotein III - house cricket (fragment)
 C:Species: Acheta domestica (house cricket)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 18-Jun-1994
 C:Accession: A60295
 R:Schacter, J.M.; Kanner, M.F.; Ziegler, P.; Wells, M.A.
 Insect Biochem. 20: 859-864, 1990
 A:Title: Apolipoprotein III: does it change? Formation of a low-density lipoprotein in the house
 cricket.
 A:Reference number: A60295
 A:Accession: A60295
 A:Molecule type: protein
 A:Residues: 1-29 <BB>
 C:Comment: This protein, a small, water-soluble apolipoprotein, is thought to increase
 is a residual pathway for flight-related lipid transport.
 C:Keywords: hemolymph; lipid transport

 Query Match 24.3% Score 26; Db 2; Length 20;
 Best Local Similarity 45.5% Pred. No. 7.5e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

 QY 9 GAGGAGGAG 19
 L L L L L L L
 Lb 3 GAGGAGGAG 13

 RESULT 7
 A55244
 tubulin alpha-1 chain - Paramesidium tetraurelia (fragment)
 C:Species: Paramesidium tetraurelia
 C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 07-Dec-1999
 C:Accession: A55244
 R:Schacter, V.; Levilliers, N.; Schmitter, J.M.; Le Goff, J.P.; Rossier, J.; Adolphe,
 Science 266: 1688-1691, 1994
 A:Title: Polyglycylation of tubulin: a posttranslational modification in axonemal microtubules
 A:Reference number: A55244; MUID:95084156
 A:Accession: A55244
 A:Status: preliminary
 C:Comment: This protein is a small, water-soluble apolipoprotein, is thought to increase
 is a residual pathway for flight-related lipid transport.
 C:Keywords: hemolymph; lipid transport
 C:Genetic code: SGG5
 C:Superfamily: tubulin

Best Local Similarity 66.7%; Pred. No. 8.5e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PPEGAQ 6
LB 5 PFQGNQ 10

Search completed: May 1, 2001, 11:26:21
Job time: 126 sec

Query Match 23.8%; Score 25.5; DB 2; Length 19;
Best Local Similarity 53.8%; Pred. No. 8.5e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 3; Gaps 1;
25 7 GLETRGVAFNVKF 19
17 1 1 1 1 1
LB 7 GD NWGVNVKF 16
RESULT 14
PT0725
Title: Receptor beta chain V-D-J region (140-21) - mouse (fragment)
Species: Mus musculus (house mouse)
Date: 17-Jul-1992 #sequence_revision 17 Jul 1992 #text_change 10-May-1997
Accession: PT0725
RefSeq: A.1
J. Exp. Med. 174, 115-124, 1991
Title: Junctional sequences of total T cell receptor beta chains have few N regions.
Reference number: PT0509; MUID:91277601
Accession: PT0725
Status: translation not shown
Molecule type: RNA
Accession: 18-PEP
Experimental source: newborn thymus, strain BALB/c
Keywords: T-cell receptor

Query Match 23.4%; Score 25; DB 2; Length 8;
Best Local Similarity 62.5%; Pred. No. 2e+05;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 5 AGRDGGK 12
1 1 1 1
LB 1 ASDNLLG 8
RESULT 15
PT0763
Title: Receptor beta chain (P15) - mouse (fragment)
Species: Mus musculus (house mouse)
Date: 17-Jul-1992 #sequence_revision 17 Jul 1992 #text_change 05-May-1999
Accession: PH0763
RefSeq: A.1; Bomer, P.; Widmann, C.; Paulilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1471-1483, 1991
Title: T cell receptor genes in a series of class I major histocompatibility complex allelic exclusion and antigen-specific repertoire.
Reference number: PH0763; MUID:92078846
Accession: PH0763
Molecule type: mRNA
Accession: 1-16-CAS
Reference: EMBL:K63957, NID:450933, FID:CAA43247.1, FID:450934
Experimental source: T lymphocyte
Keywords: T cell receptor

Query Match 23.4%; Score 25; DB 2; Length 16;
Best Local Similarity 62.5%; Pred. No. 2e+05;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 5 AGRDGGK 12
1 1 1 1
LB 1 ASDNLLG 8
RESULT 15
PT0763
Title: Receptor beta chain (P15) - mouse (fragment)
Species: Mus musculus (house mouse)
Date: 17-Jul-1992 #sequence_revision 17 Jul 1992 #text_change 05-May-1999
Accession: PH0763
RefSeq: A.1; Bomer, P.; Widmann, C.; Paulilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1471-1483, 1991
Title: T cell receptor genes in a series of class I major histocompatibility complex allelic exclusion and antigen-specific repertoire.
Reference number: PH0763; MUID:92078846
Accession: PH0763
Molecule type: mRNA
Accession: 1-16-CAS
Reference: EMBL:K63957, NID:450933, FID:CAA43247.1, FID:450934
Experimental source: T lymphocyte
Keywords: T cell receptor

Query Match 23.4%; Score 25; DB 2; Length 16;
Best Local Similarity 62.5%; Pred. No. 2e+05;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 5 AGRDGGK 12
1 1 1 1
LB 1 ASDNLLG 8
RESULT 15
PT0763
Title: Receptor beta chain (P15) - mouse (fragment)
Species: Mus musculus (house mouse)
Date: 17-Jul-1992 #sequence_revision 17 Jul 1992 #text_change 05-May-1999
Accession: PH0763
RefSeq: A.1; Bomer, P.; Widmann, C.; Paulilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1471-1483, 1991
Title: T cell receptor genes in a series of class I major histocompatibility complex allelic exclusion and antigen-specific repertoire.
Reference number: PH0763; MUID:92078846
Accession: PH0763
Molecule type: mRNA
Accession: 1-16-CAS
Reference: EMBL:K63957, NID:450933, FID:CAA43247.1, FID:450934
Experimental source: T lymphocyte
Keywords: T cell receptor


```

X NBI_LAXID 6494;
RN 1;
RP SEQUENCE;
RQ 115066 Vennom;
EX MEDLINE 96266175; PubMed=8479648;
FA Fainzilber M., Lander J.C., de J. Scherer K.C., Li K.W., Yu Z.,
EZ Fainzilber A., Bercovich W.M., et al.
R1 "A novel hydrophobic omega-conotoxin blocks molluscan dihydropyridine-
R2 sensitive calcium channels."
R3 Biochemistry 35:8748-8752(1996).
R4 1. FURTELLI M., OMIDA GH F, THE ACT AT PRESYNAPTIC MEMBRANES, THEY BIND
R5 AND BLOCK THE CALCIUM CHANNELS, SPECIFICALLY ACTS ON R TYPE
R6 CHANNELS, IN BLOCKING MOLLUSCAN DIIHYDROPYRIDINE SENSITIVE CALCIUM
R7 CHANNELS.
R8 1. MASS SPECTROMETRY, MW=2922.24, METHOD DIPOPTIC-SPRAY
R9 Presynaptic neurotoxin, Calcium channel inhibitor; Vennom.
R10 DISULFID 1 16 BY SIMILARITY.
R11 DISULFID 9 20 BY SIMILARITY.
R12 DISULFID 15 24 BY SIMILARITY.
R13 SEQUENCE 26 AA; 2640 MW; 3AFER2198666294 CRC64;

Query Match 22.4%; Score 24; DB 1; Length 26;
Best Local Similarity 40.0%; Pred. No. 1,1e+03;
Matches 4; Conservative 0; Mismatches 0; Gaps 0;

25 8 DPTGGAFAV 17
26 1 1 1 1 1
27 14 DPTGGAFAV 24

RESULT 1
EX_A_CONST 1
AC 128851;
R1 01-DEC-1992 (Rel. 24, Created)
R2 01-DEC-1992 (Rel. 24, Last sequence update)
R3 01-NOV-1997 (Rel. 35, Last annotation update)
R4 OMEGA-CONOTOXIN SVIR.
R5 Conus striatus (Striated cone).
R6 Eukaryotic; Mollusca; Gastropoda; Caudofoveata;
R7 Neogastropoda; Conoida; Conidae; Conus.
R8 NBI_LAXID 6494;
RN 11;
RP SEQUENCE AND SYNTHESIS.
RQ 115066 Vennom;
EX MEDLINE 96094172; PubMed=1609774;
FA Kamitani Y., Zolotarev S.G., Katsuki L., Hammerland L.D., Yoshikawa S.,
EZ Gray W.R., Krishnapati P., Ramachandran J., Millhauch C.C., Oliveira R.M.,
R1 Chu Z.
R2 "Novel alpha and omega-conotoxins from Conus striatus venom."
R3 Biochemistry 31:9919-9926(1992).
R4 1. Mol. Biol. 263:297-310(1996).
R5 1. FURTELLI M., OMIDA GH F, THE ACT AT PRESYNAPTIC MEMBRANES, THEY BIND
R6 AND BLOCK THE CALCIUM CHANNELS.
R7 PIR: C44379; C44379.
R8 PIR: 1MWJ; 12-AUG-97.
R9 Presynaptic neurotoxin, Calcium channel inhibitor; Vennom.
R10 Amino acid; 3D-structure.
R11 DISULFID 1 16
R12 DISULFID 8 20
R13 DISULFID 15 26
R14 M-VALUES 26 26 AMIDATION.
R15 SEQUENCE 26 AA; 2746 MW; EAD6C3F38341947F CRC64;

```

```

Query Match 22.4%; Score 24; DB 1; Length 26;
Best Local Similarity 27.3%; Pred. No. 1,1e+04;
Matches 3; Conservative 4; Mismatches 0; Gaps 0;

QY 6 QSGGTRVAFN 16
25 1 1 1 1 1
26 4 QSGGTRVAFN 14

RESULT 10
EX_A_CONST 1
AC 128851;
R1 21-JUL-1986 (Rel. 01, Created)
R2 21-JUL-1986 (Rel. 01, Last sequence update)
R3 01-FEB-1996 (Rel. 33, Last annotation update)
R4 MEDICININ 111 (IC50 IN AIX-111).
R5 Anemonia sulcata (Snake looks sea anemone).
R6 Eukaryotic; Metazoa; Cnidaria; Anthozoa; Scyphozoa; Actinoptera;
R7 Rynandidae; Actinidae; Anemonia.
R8 NBI_LAXID 6108;
RN 111;
RP SEQUENCE.
RQ MEDLINE 78984776; PubMed 24811;
FA Martinov S., Kopylov P., Schmitt H., Lachnitski M.;
R1 "Toxin 111 from Anemonia sulcata: Primary structure."
R2 FEBS Lett. 94:247-252(1977).
R3 121;
R4 PRELIMINARY SEQUENCE.
R5 MEDLINE 78044787; PubMed=21843;
R6 Beress L., Wundtner G., Wachter E.;
R7 "Amino acid sequence of toxin 111 from Anemonia sulcata."
R8 Hoppe-Seyler's Z. Physiol. Chem. 358:985-988(1977).
R9 141;
R10 STRUCTURE BY NMR.
R11 MEDLINE 93341891; PubMed 8182051;
R12 Norton R.S., Gross K., Branch-Maksvytis V., Wachter E.;
R13 "1H-NMR study of the solution properties and secondary structure of
R14 neurotoxin 111 from the sea anemone Anemonia sulcata."
R15 Blochom. J. 293:545-551(1993).
R16 141;
R17 STRUCTURE BY NMR.
R18 MEDLINE 95244415; PubMed=727358;
R19 Manoleras N., Norton R.S.;
R20 "Three-dimensional structure in solution of neurotoxin 111 from the
R21 sea anemone Anemonia sulcata."
R22 Biochemistry 33:11051-11061(1994).
R23 1. FURTELLI M., OMIDA GH F, THE ACT AT PRESYNAPTIC MEMBRANES, THEY BIND
R24 1-1. SIMILARITY: BELONGS TO THE ANEMONIN SHEET TOXINS FAMILY.
R25 PIR: A01798; TZA43.
R26 PIR: 1AMS; 31-AUG-94.
R27 Toxin; Sodium channel inhibitor; 3D-structure.
R28 DISULFID 4 17 PROBABLE.
R29 DISULFID 4 11 PROBABLE.
R30 DISULFID 6 22 PROBABLE.
R31 DISULFID 22 27 CS: 90 (IN REF. 2).
R32 SEQUENCE 27 AA; 2548 MW; AAT2261FNF4A7A CRC64;

Query Match 22.4%; Score 24; DB 1; Length 27;
Best Local Similarity 66.7%; Pred. No. 1,1e+04;
Matches 4; Conservative 0; Mismatches 2; Gaps 0;

QY 2 DECAQS 7
25 1 1 1 1 1
26 19 DECAQS 24

RESULT 11
EX_A_CONST 1
AC 128851;
R1 01-FEB-1991 (Rel. 17, Created)

```


DK PIR: A0286; SMMR.
 KW Metal-binding: Metal-thiolate cluster; Chelation: Copper.
 SC SEQUENCE 25 AA: 223 MW: 2426794.17
 or send an email to license@sib.ch).
 CC
 DR EMBL: X03009; CAA26794.1; -
 DR EMBL: M59836; AAA43594.1; -
 DR EMBL: M27709; AAA43595.1; -
 DR PIR: A24641; SMMR.
 KW Metal binding; Metal thiolate cluster; Chelation: Copper.
 FT INIT_MET 0 0
 FT METAL 3 4 COPPER.
 FT METAL 5 5
 FT METAL 11 11 COPPER.
 FT METAL 14 14 COPPER.
 FT METAL 17 17 COPPER.
 FT METAL 19 19 COPPER.
 FT METAL 22 22 COPPER.
 FT CONFLICT 21 21 N 1 (IN REF. 2).
 CC SEQUENCE 25 AA: 2234 MW: 25636046142295 GR94;
 Query Match 21.0%; Score 22.5; DR 1; Length 25;
 Best Local Similarity 55.6%; Pred. No. 1.8e+03;
 Matches 5; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 QY 7 GD-GRGVA 14
 L 1 1 1 1
 DB 1 GDCGSCAS 9
 Search completed: May 1, 2001, 11:25:51
 Job time: 101 sec
 RESULTS 15
 MI_NHEDR
 ID MI_NHEDR STARWARE PIR 25 AA.
 AC P2287;
 DI 21-JUL-1996 (Ref. 01; Created)
 DI 21-JUL-1986 (Ref. 01; Last sequence update)
 DI 15-JUL-1999 (Ref. 48; Last annotation update)
 DE METALLOTHIONEIN (MT).
 GN MT.
 OS Neurospora crassa.
 SC Escherichia: Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 SC Sordariales; Sordariaceae; Neurospora.
 CX NBL_54810-5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 91112714; PubMed 2942441;
 RA Munter K., German U.A., Lerch K.;
 RI "Isolation and structural organization of the Neurospora crassa
 K1 copper metallothionein gene";
 RI EMBL 3: 4:2465-2668(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 91112714; PubMed 2148862;
 RA Beyer F.M., Jacobs F.A., Brausseau P.;
 RI "Expression of a Neurospora crassa metallothionein and its variants
 K1 in Escherichia coli";
 RI A47. Environ. Microbiol. 56:2748-2754(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 92029998; PubMed 2959528;
 RA Munter K., German U.A., Lerch K.;
 RI "Isolation and regulation of expression of the Neurospora crassa
 K1 copper metallothionein gene";
 RI Experientia Suppl. 52:194-199(1997).
 RN [4]
 RP SEQUENCE.
 RX MEDLINE 9014244; PubMed 6444697;
 RA Lerch K.;
 RI "Upper metallothionein, a copper-binding protein from Neurospora
 K1 crassa";
 RI Nature 284: 368-370(1980).
 RN [5]
 RP SEQUENCE BY NMR.
 RX MEDLINE 8938287; PubMed 2525920;
 RA Malikyan J.A., Lerch K., Amilage J.M.;
 RI "Proton NMR studies of a metallothionein from Neurospora crassa:
 K1 sequence-specific assignments by NMR measurements in the rotating
 frame";
 RI Biochemistry 28:2941-2945(1989).
 CC -1- INDICATION: BY COPPER.
 CC -1- MISCELLANEOUS: THE GIVEN CYSTEINES BIND SIX COPPER (SUPEROXIDE) IONS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 8 IN METALLOTHIONEIN SUPERFAMILY.
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license from the EMBL Outstation at the European
 Bioinformatics Institute.

Copyright (c) 1994-2000 Compugen Ltd.

OM protein protein search using SW model

Run on: May 1, 2001, 11:27:05, Search time 40.82 seconds (without alignments)

45,941 Million cell updates/sec

File: us-09-341-829a_5_copy_164_179

Ported source: 62

Sequences: 1 KARDITIKIRVSEQR 16

Scoring table: 45 SUM62

Gapop 10.0, Gapex 0.1

Sequences: 34700 seqs, 11720716 residues

Total number of hits satisfying chosen parameters: 10868

Minimum db seq length: 6

Maximum db seq length: 60

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 4% summaries

Database:

SPIRITMIL15:

1: sp archaea:

2: sp bacteria:

3: sp fungi:

4: sp human:

5: sp invertebrates:

6: sp mammals:

7: sp other:

8: sp other:

9: sp other:

10: sp plant:

11: sp protists:

12: sp unclassified:

13: sp vertebrates:

14: sp virus:

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

SUMMARIES

| Pos. # | No. | Score | Match | Length | ID | Description |
|--------|-----|-------|-------|--------|-------|-------------------|
| 1 | 1 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 2 | 2 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 3 | 3 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 4 | 4 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 5 | 5 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 6 | 6 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 7 | 7 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 8 | 8 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 9 | 9 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 10 | 10 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 11 | 11 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 12 | 12 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 13 | 13 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 14 | 14 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 15 | 15 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 16 | 16 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 17 | 17 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 18 | 18 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |
| 19 | 19 | 26 | 41.7 | 21 | Q9X17 | Q9X177 prochloroc |

ALIGNMENTS

RESULT 1

Q9X17

ID Q9X17

AC Q9X17

DI 01-NOV-1999 (TEMBLREL, 12, Created)

DT 01-NOV-1999 (TEMBLREL, 12, Last sequence update)

DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT)

GN PETD

OS Prochlorococcus sp.

OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcus

OX NCBI TaxID: 1220

EN 11

RP SEQUENCE FROM N.A.

RA Urbach E., Chisholm S.W.

PT "Genetic diversity in Prochlorococcus populations: flow cytometrically

RT sorted from the Sargasso Sea and Gulf Stream."

RL Limnol. oceanogr., 43:1615-1640(1998)

DR EMBL: AF070208; AAU23258.1

ET Nucleic

SO SEQUENCE

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

PRELIMINARY:

```

OS Homo rhinovirus
CA Virus: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
AC Rhinovirus.
OX NCBI_TaxID:12133;
RN [1]
RA Clark G.;
KL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL109964; CAB91668.1;
FT NON_CODING
SQ SEQUENCE 28 AA: 2296 MW: 405EA521900B399 CRC64:

Query Match: 34.1%; Score 29; DB 6; Length 27;
Best local Similarity 66.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 KPRKVSQQR 16
DB 1 1 1 1 1
5 KPRKVSQQR 17

RESULT 4
Q91K26 Q91K26 PRELIMINARY: PRT: 27 AA.
AC Q91K26
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE ADRENOMEDULLIN.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID:9824;
RN [1]
RA
SQ SEQUENCE FROM N.A.
MEDLINE 96157714; PubMed 8576091;
FA Eishi Y., Yamada K., Endaka K., Sawamoto M., Matsuo H., Eto F.;
"Isolation and characterization of immunoreactive adrenomedullin in
porcine tissue, and isolation of adrenomedullin [26-52] and
adrenomedullin [34-52] from porcine duodenum.";
KL J. Biochem. 118:765-770(1995).
DR INTERPRO: IPR001710;
PF PNTS: PRO001; ADRENOMEDULLIN.
SQ SEQUENCE 27 AA: 4363 MW: B8D67FA18DBB4D90 CRC64:

Query Match: 34.1%; Score 29; DB 6; Length 27;
Best local Similarity 40.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 KPRKVSQQR 15
DB 1 1 1 1 1
11 KPRKVSQQR 25

RESULT 4
Q91K26 Q91K26 PRELIMINARY: PRT: 28 AA.
AC Q91K26
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE G11118M15.1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RA SEQUENCE FROM N.A.
RA Clark G.;
KL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL109964; CAB91668.1;
FT NON_CODING
SQ SEQUENCE 28 AA: 2296 MW: 405EA521900B399 CRC64:

Query Match: 34.1%; Score 28; DB 4; Length 28;
Best local Similarity 38.5%; Pred. No. 3.6e+02;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 4 LKPKKVSQQR 16
DB 1 1 1 1 1
7 DGRSPKVSQQR 19

RESULT 5
Q91K26 Q91K26 PRELIMINARY: PRT: 25 AA.
AC Q91K26
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE IIA CLASS II-ASSOCIATED INVARIANT CHAIN 11.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RA
SQ SEQUENCE.
MEDLINE 92375195; PubMed 1380674;
FA Chicz R.M., Urban R.G., Lane W.S., Gorra J.C., Stern L.J.,
Vignali D.A., Strominger J.L.;
"Predominant naturally processed peptides bound to HLA DR1 are derived
from MHC-related molecules and are heterogeneous in size.";
KL Nature 358:764-768(1992).
KW MHC.
SQ SEQUENCE 25 AA: 2793 MW: 2F80109276A802F4 CRC64:

Query Match: 31.7%; Score 26; DB 7; Length 25;
Best local Similarity 50.0%; Pred. No. 7.3e+02;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 5 LKPKKVSQQR 16
DB 1 1 1 1 1
1 LKPKKVSQQR 12

RESULT 6
Q91K26 Q91K26 PRELIMINARY: PRT: 27 AA.
AC Q91K26
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE POLYPROTEIN (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RA SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.;
"The genetic heterogeneity of hypervariable region 1 of the viral
RNA and the sensitivity of hepatitis C virus to interferon alpha
therapy.";
KL Sub.itted (MAY 1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF166797; AAC5365.1;
DR INTERPRO: IPR002531;

```

DR PFAM: PF01560; HCV_NSI: 1.
 KW Polypeptide.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2945 MW; 911B4189A752B140 CR664;

Query Match: 41.7%; Score 26; Ids 14; Length 27;
 Best Local Similarity: 83.9%; Pred. No. 7.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADJRT 7
 11111
 Db 10 ADVRT 15

RESULT 7
 Q90167 PRELIMINARY; PFI: 27 AA.
 AC Q90167; 01 MAY 2000 (TEMBLrel. 14, Created)
 DT 01 MAY 2000 (TEMBLrel. 14, Last sequence update)
 DI 01 JUN 2000 (TEMBLrel. 14, Last annotation update)
 DE POLYPEPTIDE (FRAGMENT).
 OS Hepatitis C virus.
 OC VIRUSES; ssRNA positive-strand viruses, no RNA stage; Flaviviridae;
 OX NCBI_TaxID:11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sandres K., Dubois M., Pasquier C., Lopez J.,
 RT "The genetic heterogeneity of hypervariable region 1 of the viral
 R1 genome and the sensitivity of hepatitis C virus to interferon alpha
 R1 therapy.";
 RL Submitted (Jul. 1999) to the EMBL/GenBank/Tran databases.
 DR EMBL: AF166710; AAC52468.1;
 DR INTERPRO: IPR002531;
 DR PFAM: PF01560; HCV_NSI: 1.
 KW Polypeptide.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2945 MW; 911B4189A752B140 CR664;

Query Match: 41.7%; Score 26; Ids 14; Length 27;
 Best Local Similarity: 83.9%; Pred. No. 7.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADJRT 7
 11111
 Db 10 ADVRT 15

RESULT 8
 Q90166 PRELIMINARY; PFI: 27 AA.
 AC Q90166; 01 MAY 2000 (TEMBLrel. 14, Created)
 DT 01 MAY 2000 (TEMBLrel. 14, Last sequence update)
 DI 01 JUN 2000 (TEMBLrel. 14, Last annotation update)
 DE POLYPEPTIDE (FRAGMENT).
 OS Hepatitis C virus.
 OC VIRUSES; ssRNA positive-strand viruses, no RNA stage; Flaviviridae;
 OX NCBI_TaxID:11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sandres K., Dubois M., Pasquier C., Lopez J.,
 RT "The genetic heterogeneity of hypervariable region 1 of the viral
 R1 genome and the sensitivity of hepatitis C virus to interferon alpha
 R1 therapy.";
 RL Submitted (Jul. 1999) to the EMBL/GenBank/Tran databases.
 DR EMBL: AF166710; AAC52468.1;

DR INTERPRO: IPR002531;
 DR PFAM: PF01560; HCV_NSI: 1.
 KW Polypeptide.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2945 MW; 911B4189A752B140 CR664;

Query Match: 41.7%; Score 26; Ids 14; Length 27;
 Best Local Similarity: 83.9%; Pred. No. 7.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADJRT 7
 11111
 Db 10 ADVRT 15

RESULT 9
 Q90167 PRELIMINARY; PFI: 27 AA.
 AC Q90167; 01 MAY 2000 (TEMBLrel. 14, Created)
 DT 01 MAY 2000 (TEMBLrel. 14, Last sequence update)
 DI 01 JUN 2000 (TEMBLrel. 14, Last annotation update)
 DE POLYPEPTIDE (FRAGMENT).
 OS Hepatitis C virus.
 OC VIRUSES; ssRNA positive-strand viruses, no RNA stage; Flaviviridae;
 OX NCBI_TaxID:11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sandres K., Dubois M., Pasquier C., Lopez J.,
 RT "The genetic heterogeneity of hypervariable region 1 of the viral
 R1 genome and the sensitivity of hepatitis C virus to interferon alpha
 R1 therapy.";
 RL Submitted (Jul. 1999) to the EMBL/GenBank/Tran databases.
 DR EMBL: AF166710; AAC52468.1;
 DR INTERPRO: IPR002531;
 DR PFAM: PF01560; HCV_NSI: 1.
 KW Polypeptide.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2945 MW; 911B4189A752B140 CR664;

Query Match: 41.7%; Score 26; Ids 14; Length 27;
 Best Local Similarity: 83.9%; Pred. No. 7.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ADJRT 7
 11111
 Db 10 ADVRT 15

RESULT 10
 Q90166 PRELIMINARY; PFI: 27 AA.
 AC Q90166; 01 MAY 2000 (TEMBLrel. 14, Created)
 DT 01 MAY 2000 (TEMBLrel. 14, Last sequence update)
 DI 01 JUN 2000 (TEMBLrel. 14, Last annotation update)
 DE POLYPEPTIDE (FRAGMENT).
 OS Hepatitis C virus.
 OC VIRUSES; ssRNA positive-strand viruses, no RNA stage; Flaviviridae;
 OX NCBI_TaxID:11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sandres K., Dubois M., Pasquier C., Lopez J.,
 RT "The genetic heterogeneity of hypervariable region 1 of the viral
 R1 genome and the sensitivity of hepatitis C virus to interferon alpha
 R1 therapy.";
 RL Submitted (Jul. 1999) to the EMBL/GenBank/Tran databases.

DR EMBL: AF166714, AA052371.1;
 DR INTERPRO: IPR002541;
 DR PFAM: PF01560; HCV_NSI; 1;
 KW Polyprotein.
 FT N_N_TER 1 1
 FT N_N_TER 27 27
 SQ SEQUENCE 27 AA; 2945 MW; 911B4D89A752813C CRC64;

Query Match 41.7%; Score 26; DB 14; Length 27;
 Best Local Similarity 84.9%; Pred. No. 7.9e-02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ARDRT 7
 III:II

DB 10 ARDVRT 15

RESULT 11

Q9Q165 PRELIMINARY: PRT: 27 AA.

AC Q9Q165; 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DI 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE POLYPROTEIN (FRAGMENT);
 OS Hepatitis C virus.

CC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OX Hepacivirus.
 OX NCBI_TaxID=11103;

RN [1]

RP SEQUENCE FROM N.A.

RA Sandres K., Dubois M., Pasquier C., Izopet J.;

ET "The genetic heterogeneity of hypervariable region 1 of the viral genome and the sensitivity of hepatitis C virus to interferon alpha therapy.";

FI 246771.1 (Apr. 1999); to the EMBL/GenBank/DBJ databases.

DR EMBL: AF166714, AA052371.1;
 DR INTERPRO: IPR002541;
 DR PFAM: PF01560; HCV_NSI; 1;

KW Polyprotein.
 FT N_N_TER 1 1
 FT N_N_TER 27 27

SQ SEQUENCE 27 AA; 2945 MW; 911B4D89A752813C CRC64;

Query Match 41.7%; Score 26; DB 14; Length 27;
 Best Local Similarity 84.9%; Pred. No. 7.9e-02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ARDRT 7
 III:II

DB 10 ARDVRT 15

RESULT 12

Q9Q164 PRELIMINARY: PRT: 27 AA.

AC Q9Q164; 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DI 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE POLYPROTEIN (FRAGMENT);

OS Hepatitis C virus.

CC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OX Hepacivirus.

OX NCBI_TaxID=11103;

RN [1]

RP SEQUENCE FROM N.A.

RA Sandres K., Dubois M., Pasquier C., Izopet J.;

ET "The genetic heterogeneity of hypervariable region 1 of the viral genome and the sensitivity of hepatitis C virus to interferon alpha therapy.";

KI Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF166714, AA052371.1;
 DR INTERPRO: IPR002541;
 DR PFAM: PF01560; HCV_NSI; 1;
 KW Polyprotein.
 FT N_N_TER 1 1
 FT N_N_TER 27 27
 SQ SEQUENCE 27 AA; 2935 MW; 911B4D89A752813C CRC64;

Query Match 41.7%; Score 26; DB 14; Length 27;
 Best Local Similarity 83.9%; Pred. No. 7.9e-02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ARDRT 7
 III:II

DB 10 ARDVRT 15

RESULT 13

Q9Q163 PRELIMINARY: PRT: 27 AA.

AC Q9Q163; 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DI 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE POLYPROTEIN (FRAGMENT);

OS Hepatitis C virus.

CC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OX Hepacivirus.

OX NCBI_TaxID=11103;

RN [1]

RP SEQUENCE FROM N.A.

RA Sandres K., Dubois M., Pasquier C., Izopet J.;

ET "The genetic heterogeneity of hypervariable region 1 of the viral genome and the sensitivity of hepatitis C virus to interferon alpha therapy.";

KI Submitted (Jul-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF166714, AA052372.1;
 DR PFAM: PF01560; HCV_NSI; 1;

KW Polyprotein.
 FT N_N_TER 1 1
 FT N_N_TER 27 27

SQ SEQUENCE 27 AA; 2943 MW; 911B4D89A752813C CRC64;

Query Match 41.7%; Score 26; DB 14; Length 27;
 Best Local Similarity 84.9%; Pred. No. 7.9e-02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ARDRT 7
 III:II

DB 10 ARDVRT 15

RESULT 14

Q9Q162 PRELIMINARY: PRT: 27 AA.

AC Q9Q162; 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DI 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE POLYPROTEIN (FRAGMENT);

OS Hepatitis C virus.

CC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OX Hepacivirus.

OX NCBI_TaxID=11103;

RN [1]

RP SEQUENCE FROM N.A.

RA Sandres K., Dubois M., Pasquier C., Izopet J.;

ET "The genetic heterogeneity of hypervariable region 1 of the viral genome and the sensitivity of hepatitis C virus to interferon alpha therapy.";

KI Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

LR EMBL: AF166216; AAI52474.1;
 LR INTERPRO: IPR002541;
 LR PFAM: PF01540; HCV_NSI; 1;
 KW Polyprotein.
 FT N N TER 1
 FT N N TER 27
 SQ SEQUENCE 27 AA: 2945 MW: 91184189A7528130 CRC64:

Query Match 41.7% Score 26; DB 14; Length 27;
 Best Local Similarity 83.0% Prod. No. 7.9e+02;

Matches 5; Conservatve 1; Mismatches 0; Indels 0; Gaps 0;

27 2 APLLEK 7

10 10 ADYK 15

PREMIL 1

Seq161

10 09061 PRELIMINARY; PRT; 27 AA.

AA 09061;

11 01 MAY 2000 (1:EMBLrel. 14, Created)

11 01 MAY 2000 (1:EMBLrel. 14, Last sequence update)

11 01 JUN 2000 (1:EMBLrel. 14, Last annotation update)

11 01 POLYPROTEIN (FRAGMENT)

11 01 Hepatitis C virus

11 01 Virusess: RNA positive-strand viruses, no RNA stage: Flaviviridae;

11 01 Hepatitis C virus

11 01 NMR 1x10 11100

11 11

11 11 SEQUENCE FROM N.A.

11 11 Sandroes K., Indels M., Pasquier C., Lopez J.

11 11 "The genetic heterogeneity of hypervariable region 1 of the viral

11 11 genome and the sensitivity of hepatitis C virus to interferon alpha

11 11 therapy."

11 11 Submitted (JUL 1999) to the EMBL/Genbank/DBJ databases.

11 11 EMBL: AF166216; AAI52474.1;

11 11 INTERPRO: IPR002541;

11 11 PFAM: PF01540; HCV_NSI; 1;

11 11 Polyprotein.

11 11 N N TER 1

11 11 N N TER 27

11 11 SEQUENCE 27 AA: 2945 MW: 91184189A7528130 CRC64:

Query Match

Best Local Similarity 83.0% Prod. No. 7.9e+02;
 Matches 5; Conservatve 1; Mismatches 0; Indels 0; Gaps 0;

27 2 APLLEK 7

10 10 ADYK 15

Search completed: May 1, 2001, 11:27:05
 Database: 145, size

OS Hepatitis C virus.
 OC Viruses; ssRNA positive strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 CX NCBI_TaxID=11103;
 GN [1]
 RP SEQUENCE FROM N.A.
 RA Sandrois K., Dubois M., Pasquier C., Izopet J.;
 RI "the genetic heterogeneity of hypervariable region 1 of the viral
 RT genome and the sensitivity of hepatitis C virus to interferon alpha
 RT therapy";
 RL Submitted (01-1999) to the EMBL/GenBank/DBJ databases.
 LF EMBL: AF166829; AA052484.1;
 DB INTERPRO: IPR002531;
 DR PFAM: PF01560; HCV_NSL; 1.
 KW Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2757 MW; 61AF606C4295BE22 CRC64;

Query Match. 29.9%, Score 32, DB 14, Length 27;
 Best Local Similarity 43.8%, Pred. No. 1.7e-02;
 Matches 7; Conservative 4; Mismatches 4; Indels 2; Gaps 1;
 QY 4 GAAGGCGGVA--FNV 17
 11 1111111
 DB 7 GSEAHGARGLASLFNM 22

RESULT 4
 Q9QH04
 ID Q9QH04 PRELIMINARY; PRI: 27 AA.
 DI 01-MAY-2000 (TEMBLrel. 13, Created)
 DI 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DI 01-JUN-2000 (TEMBLrel. 14, Last annotation update)
 DE POLYPROTEIN (FRAGMENT).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 CX NCBI_TaxID=11103;
 GN [1]
 RP SEQUENCE FROM N.A.
 RA Sandrois K., Dubois M., Pasquier C., Izopet J.;
 RI "the genetic heterogeneity of hypervariable region 1 of the viral
 RT genome and the sensitivity of hepatitis C virus to interferon alpha
 RT therapy";
 RL Submitted (01-1999) to the EMBL/GenBank/DBJ databases.
 LF EMBL: AF166829; AA052484.1;
 DB INTERPRO: IPR002531;
 DR PFAM: PF01560; HCV_NSL; 1.
 KW Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2757 MW; 61AF606C4295BE22 CRC64;

Query Match. 29.9%, Score 32, DB 14, Length 27;
 Best Local Similarity 43.8%, Pred. No. 1.7e-02;
 Matches 7; Conservative 4; Mismatches 4; Indels 2; Gaps 1;
 QY 4 GAAGGCGGVA--FNV 17
 11 1111111
 DB 7 GSEAHGARGLASLFNM 22

RESULT 4
 Q9QH04
 ID Q9QH04 PRELIMINARY; PRI: 27 AA.
 DI 01-MAY-2000 (TEMBLrel. 13, Created)
 DI 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DI 01-JUN-2000 (TEMBLrel. 14, Last annotation update)

DE POLYPROTEIN (FRAGMENT).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 CX NCBI_TaxID=11103;
 GN [1]
 RP SEQUENCE FROM N.A.
 RA Sandrois K., Dubois M., Pasquier C., Izopet J.;
 RI "the genetic heterogeneity of hypervariable region 1 of the viral
 RT genome and the sensitivity of hepatitis C virus to interferon alpha
 RT therapy";
 RL Submitted (01-1999) to the EMBL/GenBank/DBJ databases.
 LF EMBL: AF166828; AA052486.1;
 DB INTERPRO: IPR002531;
 DR PFAM: PF01560; HCV_NSL; 1.
 KW Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2757 MW; 61AF606C4295BE22 CRC64;

Query Match. 29.9%, Score 32, DB 14, Length 27;
 Best Local Similarity 43.8%, Pred. No. 1.7e-02;
 Matches 7; Conservative 4; Mismatches 4; Indels 2; Gaps 1;
 QY 4 GAAGGCGGVA--FNV 17
 11 1111111
 DB 7 GSEAHGARGLASLFNM 22

RESULT 5
 Q9QH08
 ID Q9QH08 PRELIMINARY; PRI: 27 AA.
 DI 01-MAY-2000 (TEMBLrel. 13, Created)
 DI 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DI 01-JUN-2000 (TEMBLrel. 14, Last annotation update)
 DE POLYPROTEIN (FRAGMENT).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 CX NCBI_TaxID=11103;
 GN [1]
 RP SEQUENCE FROM N.A.
 RA Sandrois K., Dubois M., Pasquier C., Izopet J.;
 RI "the genetic heterogeneity of hypervariable region 1 of the viral
 RT genome and the sensitivity of hepatitis C virus to interferon alpha
 RT therapy";
 RL Submitted (01-1999) to the EMBL/GenBank/DBJ databases.
 LF EMBL: AF166829; AA052487.1;
 DB INTERPRO: IPR002531;
 DR PFAM: PF01560; HCV_NSL; 1.
 KW Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA; 2757 MW; 61AF606C4295BE22 CRC64;

Query Match. 29.9%, Score 32, DB 14, Length 27;
 Best Local Similarity 43.8%, Pred. No. 1.7e-02;
 Matches 7; Conservative 4; Mismatches 4; Indels 2; Gaps 1;
 QY 4 GAAGGCGGVA--FNV 17
 11 1111111
 DB 7 GSEAHGARGLASLFNM 22

RESULT 6
 Q9QH07
 ID Q9QH07 PRELIMINARY; PRI: 27 AA.
 DI 01-MAY-2000 (TEMBLrel. 13, Created)
 DI 01-MAY-2000 (TEMBLrel. 13, Last sequence update)

DI 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE POLYPROTEIN (FRAGMENT)
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive strand viruses, no DNA stage; Flaviviridae;
 OX NCBITaxID:11104;
 RN 111
 RP SEQUENCE FROM N.A.
 RA Sandres K., Dubois M., Pasquier C., Lopez J.J.
 RI "The genetic heterogeneity of hypervariable region 1 of the viral
 RI genome and the sensitivity of hepatitis C virus to interferon alpha
 RI therapy.";
 RL Submitted (JUL 1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF166837; AAU52496.1;
 DR INTERPRO: IPR002531;
 DR PFAM: PF01560; HCV_NSI: 1;
 KW Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA: 2757 MW: 61AF606C4295DE22 CR664;

Query Match 29.9% Score 32; DB 14; Length 27;

Best Local Similarity 43.8% Pred. No. 1,76,02;

Matches 7; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 4 CAQGLGGRGVA FNV 17

I: 1 1 1 1 1 1 1

DB 7 GSEAHGARGLASLFNM 22

RESULT 7

Q9QHT9

AC Q9QHT9 PRELIMINARY; PRI: 27 AA.

DI 01-MAY-2000 (TrEMBLrel. 14, Created)

DE 01-MAY-2000 (TrEMBLrel. 14, Last sequence update)

DI 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

DE POLYPROTEIN (FRAGMENT)

OS Hepatitis C virus.

OC Viruses; ssRNA positive strand viruses, no DNA stage; Flaviviridae;

OX NCBITaxID:11104;

RN 111

RP SEQUENCE FROM N.A.

RA Sandres K., Dubois M., Pasquier C., Lopez J.J.

RI "The genetic heterogeneity of hypervariable region 1 of the viral

RI genome and the sensitivity of hepatitis C virus to interferon alpha

RI therapy.";

RL Submitted (JUL 1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF166837; AAU52496.1;

DR INTERPRO: IPR002531;

DR PFAM: PF01560; HCV_NSI: 1;

KW Polyprotein.

FT NON_TER 1 1

FT NON_TER 27 27

SQ SEQUENCE 27 AA: 2757 MW: 61AF606C4295DE22 CR664;

Query Match 29.9% Score 32; DB 14; Length 27;

Best Local Similarity 43.8% Pred. No. 1,76,02;

Matches 7; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 4 CAQGLGGRGVA FNV 17

I: 1 1 1 1 1 1 1

DB 7 GSEAHGARGLASLFNM 22

RESULT 4

Q9QHT9

AC Q9QHT9 PRELIMINARY; PRI: 27 AA.

DI 01-MAY-2000 (TrEMBLrel. 14, Created)

DI 01-MAY-2000 (TrEMBLrel. 14, Last sequence update)
 DE 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE POLYPROTEIN (FRAGMENT)
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive strand viruses, no DNA stage; Flaviviridae;
 OX NCBITaxID:11104;
 RN 111
 RP SEQUENCE FROM N.A.
 RA Sandres K., Dubois M., Pasquier C., Lopez J.J.
 RI "The genetic heterogeneity of hypervariable region 1 of the viral
 RI genome and the sensitivity of hepatitis C virus to interferon alpha
 RI therapy.";
 RL Submitted (JUL 1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF166837; AAU52496.1;
 DR INTERPRO: IPR002531;
 DR PFAM: PF01560; HCV_NSI: 1;
 KW Polyprotein.
 FT NON_TER 1 1
 FT NON_TER 27 27
 SQ SEQUENCE 27 AA: 2757 MW: 61AF606C4295DE22 CR664;

Query Match 29.9% Score 32; DB 14; Length 27;

Best Local Similarity 43.8% Pred. No. 1,76,02;

Matches 7; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 4 CAQGLGGRGVA FNV 17

I: 1 1 1 1 1 1 1

DB 7 GSEAHGARGLASLFNM 22

RESULT 9

Q9QHT9

AC Q9QHT9 PRELIMINARY; PRI: 27 AA.

DI 01-MAY-2000 (TrEMBLrel. 14, Created)

DE 01-MAY-2000 (TrEMBLrel. 14, Last sequence update)

DI 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

DE POLYPROTEIN (FRAGMENT)

OS Hepatitis C virus.

OC Viruses; ssRNA positive strand viruses, no DNA stage; Flaviviridae;

OX NCBITaxID:11104;

RN 111

RP SEQUENCE FROM N.A.

RA Sandres K., Dubois M., Pasquier C., Lopez J.J.

RI "The genetic heterogeneity of hypervariable region 1 of the viral

RI genome and the sensitivity of hepatitis C virus to interferon alpha

RI therapy.";

RL Submitted (JUL 1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF166837; AAU52496.1;

DR INTERPRO: IPR002531;

DR PFAM: PF01560; HCV_NSI: 1;

KW Polyprotein.

FT NON_TER 1 1

FT NON_TER 27 27

SQ SEQUENCE 27 AA: 2757 MW: 61AF606C4295DE22 CR664;

Query Match 29.9% Score 32; DB 14; Length 27;

Best Local Similarity 43.8% Pred. No. 1,76,02;

Matches 7; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 4 CAQGLGGRGVA FNV 17

I: 1 1 1 1 1 1 1

DB 7 GSEAHGARGLASLFNM 22

RESULT 10

Q9QHT9

AC Q9QHT9 PRELIMINARY; PRI: 27 AA.

DT 01-MAY-2000 (TREMBLER, 13, Created)
DI 01-MAY-2000 (TREMBLER, 13, Last sequence update)
DI 01-JUN-2000 (TREMBLER, 14, Last annotation update)
DE POLYPROTEIN (FRAGMENT)
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
CX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.
RT "The genetic heterogeneity of hypervariable region 1 of the viral
genome and the sensitivity of hepatitis C virus to interferon alpha
therapy.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF166839; AAF52501.1;
DR INTERPRO: IPR002541;
DR PFAM: PF01560; HCV_NSI; 1.
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2757 MW; 6DAF66C4295BE22 CRC64;

Query Match 29.9% Score 32; DB 14; Length 27;
Best Local Similarity 43.9% Pred. NO; 1.7e+02;
Matches 7; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 4 GAQDNGRGVA--FNV 17
111 11111111
DB 7 GSEAHGARGLASLFNM 22

RESULT 11

Q9QHT6 PRELIMINARY; PRT; 27 AA.
AC Q9QHT6;
DI 01-MAY-2000 (TREMBLER, 13, Created)
DI 01-MAY-2000 (TREMBLER, 13, Last sequence update)
DI 01-JUN-2000 (TREMBLER, 14, Last annotation update)
DE POLYPROTEIN (FRAGMENT)
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
CX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.
RT "The genetic heterogeneity of hypervariable region 1 of the viral
genome and the sensitivity of hepatitis C virus to interferon alpha
therapy.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF166843; AAF52501.1;
DR INTERPRO: IPR002541;
DR PFAM: PF01560; HCV_NSI; 1.
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2757 MW; 6DAF66C4295BE22 CRC64;

Query Match 29.9% Score 32; DB 14; Length 27;
Best Local Similarity 43.9% Pred. NO; 1.7e+02;
Matches 7; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 4 GAQDNGRGVA--FNV 17
111 11111111
DB 7 GSEAHGARGLASLFNM 22

RESULT 12

Q9QHT5 PRELIMINARY; PRT; 27 AA.
ID Q9QHT5

AC Q9QHT5;
DI 01-MAY-2000 (TREMBLER, 13, Created)
DI 01-MAY-2000 (TREMBLER, 13, Last sequence update)
DI 01-JUN-2000 (TREMBLER, 14, Last annotation update)
DE POLYPROTEIN (FRAGMENT)
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
CX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.
RT "The genetic heterogeneity of hypervariable region 1 of the viral
genome and the sensitivity of hepatitis C virus to interferon alpha
therapy.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF166842; AAF52500.1;
DR INTERPRO: IPR002541;
DR PFAM: PF01560; HCV_NSI; 1.
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2757 MW; 6DAF66C4295BE22 CRC64;

Query Match 29.9% Score 32; DB 14; Length 27;
Best Local Similarity 43.9% Pred. NO; 1.7e+02;
Matches 7; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 4 GAQDNGRGVA--FNV 17
111 11111111
DB 7 GSEAHGARGLASLFNM 22

RESULT 13

Q9QHT4 PRELIMINARY; PRT; 27 AA.
AC Q9QHT4;
DI 01-MAY-2000 (TREMBLER, 13, Created)
DI 01-MAY-2000 (TREMBLER, 13, Last sequence update)
DI 01-JUN-2000 (TREMBLER, 14, Last annotation update)
DE POLYPROTEIN (FRAGMENT)
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
CX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sandres K., Dubois M., Pasquier C., Izopet J.
RT "The genetic heterogeneity of hypervariable region 1 of the viral
genome and the sensitivity of hepatitis C virus to interferon alpha
therapy.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF166843; AAF52501.1;
DR INTERPRO: IPR002541;
DR PFAM: PF01560; HCV_NSI; 1.
KW Polyprotein.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2757 MW; 6DAF66C4295BE22 CRC64;

Query Match 29.9% Score 32; DB 14; Length 27;
Best Local Similarity 43.9% Pred. NO; 1.7e+02;
Matches 7; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

QY 4 GAQDNGRGVA--FNV 17
111 11111111
DB 7 GSEAHGARGLASLFNM 22

RESULT 14

Q9QHT3

```

10 QUERY: PRELIMINARY: PRI: 27 AA.
AC QUERY:
D1 01 MAY 2000 (EMBLrel. 14, Created)
D1 01 MAY 2000 (EMBLrel. 14, Last sequence update)
D1 01 JUN 2000 (EMBLrel. 14, Last annotation update)
DE POLYPEPTIDE (FRAGMENT).
OS Hepatitis C virus.
OC Viruses; ssRNA positive strand viruses, no DNA stage; Flaviviridae;
OC Hepaviruses.
OX NCBI Taxid: 1114;
PN 111
PR SEQUENCE FROM N.A.
PA Sandroes K., Lubasi M., Passauer C., Lopez J.;
P1 "The genetic heterogeneity of hypervariable region 1 of the viral
P1 genome and the sensitivity of hepatitis C virus to interferon alpha
P1 therapy.";
RC Submitted (JUL 1999) to the EMBL/GenBank/DBJ databases.
RC EMBL: AF166944; AAL52502.1;
DE INTERFERON (FRAGMENT).
DE EFAM: P41560; REV_NS1; 1
KW Polypeptide.
FI NON TER 1 1
FI NON TER 27 27
SQ SEQUENCE 27 AA: 2759 MW: 31466664295RE22 CRG64;

Query Match
Best Local Similarity 29.4% Score 32; DB 14; Length 27;
Matches 7; Conserved 4; Mismatches 3; Indels 2; Gaps 1;

Q7 4 GAGAGGAGCA FNV 17
1: 1 1 1 1 1
DB 7 GSEMPACELASHNM 22

RESULT 15
Q99922
AC Q99922 PRELIMINARY: PRI: 40 AA.
D1 01 MAY 1997 (EMBLrel. 04, Created)
D1 01 MAY 1997 (EMBLrel. 04, Last sequence update)
D1 01 MAY 1999 (EMBLrel. 10, Last annotation update)
DE CYCLIN D1 (H3-GAMMA-1 PROTEIN (FRAGMENT)).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI Taxid: 9606;
PN 111
PR SEQUENCE FROM N.A.
PA MEDLINE: 94269516; PubMed 8696815;
PA Chesel M., Boissacq P.L., Brotons L.A., Smith C.M., Gerhard D.S.,
PA Kuhl W.M.;
P1 "Dysregulation of cyclin D1 by translocation into an IgH gamma switch
P1 region in two multiple myeloma cell lines.";
DE EMBL: S82568; AAL34412.1;
DE EMBL: S82568; AAL34412.1;
FI NON TER 1 1
FI NON TER 40 40
SQ SEQUENCE 40 AA: 2988 MW: 4591674992P699 CRG64;

Query Match
Best Local Similarity 29.4% Score 31.5; DB 4; Length 40;
Matches 7; Conserved 0; Mismatches 3; Indels 1; Gaps 1;

Q7 2 PEAGAGGQPR 11
1: 1 1 1 1 1
DB 7 PAGEAGGQWER 17

```

